English Language Education

Susanne M. Reiterer *Editor*

Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience

Springer

English Language Education

Volume 16

This series publishes research on the development, implementation and evaluation of educational programs for school-aged and adult learners for whom English is a second or additional language, including those who are learning academic content through the medium of English. The series has a dual focus on learners' language development and broader societal and policy-related issues, including the implications for teachers' professional development and policy support at the institutional and system level. The series seeks to engage with current issues in English language teaching (ELT) in educational institutions from a highly situated standpoint, examining theories, practices and policies with a conscious regard for historical lineages of development and local (re)contextualisation. By focusing on multiple educational contexts and adopting a comparative perspective, the series will transcend traditional geographical boundaries, thus will be relevant to both English-speaking countries and countries where English is a very much an additional, but important language for learning other content. This series will also cross disciplinary and methodological boundaries by integrating sociocultural and critical approaches with second language acquisition perspectives and drawing on both applied linguistics and educational research. In drawing together basic and applied policy-related research concerns, the series will contribute towards developing a more comprehensive, innovative and contextualized view of English language education internationally. Authors are invited to approach the Series Editor with ideas and plans for books. For more information, please contact the publishing editor, Jolanda Voogd. E-mail: jolanda.voogd@springer.com

More information about this series at http://www.springer.com/series/11558

Susanne M. Reiterer
Editor

Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience

Springer

Editor
Susanne M. Reiterer
Department of Linguistics and Centre
for Teacher Education
University of Vienna
Vienna, Wien, Austria

ISSN 2213-6967 ISSN 2213-6975 (electronic)
English Language Education
ISBN 978-3-319-91916-4 ISBN 978-3-319-91917-1 (eBook)
https://doi.org/10.1007/978-3-319-91917-1

Library of Congress Control Number: 2018950568

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedication – to my dear mother who passed away so early

Preface

Language and linguistic behaviour came into the world by biological evolution and by its speakers who shaped all the variation in language which we can scientifically observe right now. Language changes over time, space, between and within individuals. This creates an ever fluctuating "stress field" between universality and individuality. Some of these manifold extralinguistic factors, which contribute to individual variation in linguistic behaviour, concerning first – but also and specifically here – second language learning, are the focus of this book which is based on a collaborative research project on "individual variation in foreign language aptitude", originating from a seminar on second language learning aptitude, the editor had given at the University of Vienna. The project was mainly based at the University of Vienna, in cooperation with the Universities of Graz and Heidelberg.

Human communication originates from social-emotional interaction, face-to-face, in its most basic form. It comprises so many different aspects of our lives: emotional, intellectual-cognitive, physical-motion/motoric, biochemical (neurotransmitters, hormones), social-interactional, sociocultural, psychological, expressive and the like. This patchwork of aspects reflects universal biological, cultural-societal as well as individual needs and personal identities. All these diverse dimensions, which create our communicative behaviour (and ultimately our language or languages as a system), show a lot of individual variation, or individual differences (IDs) – a term more often used in psycholinguistics and psychology – giving rise to individual language learner and speaker profiles. This variation is seen when speaking and acquiring mother tongues, dialects and foreign languages alike. In their first languages, people are so different in their communicative behaviour that it can reach from mutism to logorrhoea, from poor expressive abilities to outstanding rhetoric gifts in orators, from dyslexia and agrammatism to eloquent genius in literary production, from unclear speech in articulation difficulties (e.g. developmental apraxia of speech) to hyper articulation capacities in comedians, parodists and impersonators or singers, as well as from autism to high pragmatic ability in communication talent. Speaker-based individual differences in speech, language and communication behaviour are simply enormous and all too often forgot in scientific models which try to capture the reality of speech and language behaviour mostly through written language sources, are simplified, assume averaged or ideal speakers' behaviour and often use a binary logic of truth (absence or presence of something, true or false). One example is the often quoted question related to variation in pronunciation performance of early versus later foreign language learners (sensitive period theory), e.g. "can a speaker with a late onset of L2 learning ever achieve native like proficiency in pronunciation?" The answer given or expected is either "yes or no". This also accounts for the often posed question "is it nature or nurture?" The answer could or should rather start with "well, it depends on the distribution of the phenomenon". In our case, language aptitude has long been known and observed to be normally distributed in populations (Gaussian distribution, Neufeld, 1979). A solution to under-realistic or dichotomous models or theories is to "model" language aptitude by using the normal distribution and predict theoretical outcomes in terms of percentages. According to this "model", around 5–15% of individuals (very roughly speaking, every 10th person) can attain phonetic native speaker pronunciation, based on their aptitude profiles. 70% of individuals arrive "only" at an average pronunciation proficiency in a later learned second language, because that can be predicted by their position in the distributional curve; however, those 70% (and not 100) are the ones for whom most models and theories have been developed. Those are the so-called "masses" or the "norm". For yet another 5–15%, it might be enormously difficult to learn foreign languages or the pronunciation of it, if we stick to the above example, and yet for another 2% (or, every 50th person), the ability barriers might make it even almost impossible or just very difficult. However, if we leave out 30% of the whole population, we can no longer talk about "inclusion", speaking in terms of education. Exception and rule should both be included and accommodated in the models and theories to arrive at a description of 100% of all individuals, namely, the whole continuum. In the present volume and in own previous research about the impact of different psychological and neurocognitive factors on language aptitude, we could repeatedly and clearly demonstrate that all aspects of language aptitude we investigated so far (e.g. aptitude for pronunciation, vocabulary learning and associative memory, syntactic sensitivity, even pragmatic or singing ability) are always normally distributed (Christiner & Reiterer, 2013, 2015; Dogil & Reiterer, 2009; Hu et al., 2013; Marusakova 2014; Reiterer et al., 2011; Reiterer, Hu, Sumathi, & Singh, 2013; Wucherer & Reiterer, 2016). Other than this, in our own research, we did not only find neurological or neurocognitive predictors of language aptitude (i.e. individual differences in pronunciation/speech imitation capacities as reflected by brain structure or different activation patterns, as in Reiterer et al., 2011; Hu et al., 2013; Vaquero, Rodriguez-Fornells, & Reiterer, 2017; Turker et al., this volume) but also acoustic-articulatory predictors reflected in characteristic "articulation space" patterns, which were larger for the high- versus low-aptitude individuals, as analysed by modulation spectrum analysis (e.g. Reiterer et al., 2013); phonetic predictors in vowel duration as in the "schwa" sound, which we observed to be as short in native L1 speakers as in very talented second language speakers' pronunciation samples (see Ghafoorian, this volume); and cross-domain cognitive predictors relevant for speech imitation aptitude, like general musicality, singing abilities and working memory (Christiner & Reiterer, 2013, 2015, 2016; Nardo & Reiterer, 2009); characteristic personality aspects (Rota & Reiterer, 2009; Hu et al., 2013), amongst which "openness for new experience" seems to relate to L2 aptitude; as well as gender differences differentially reflecting language aptitude for speech imitation versus grammar and vocabulary learning (Wucherer & Reiterer, 2016; Habl, this volume).

The normal distribution of a phenomenon (be it aptitude, body weight, body size, etc.) also points at a potentially underlying biological system. However, there is still paucity of research into the biological, biochemical or genetic roots of language aptitude (probably due to financial and methodological complexities and constraints), apart from a very laudable recent upstream and increased interest into the genetic foundations and hereditability of second language learning/acquisition (as opposed to first language acquisition) and language abilities in general and the individual differences thereof (e.g. see Dale, Harlaar, Haworth, & Plomin, 2010; Hayiou, Dale, & Plomin, 2012; Dediu, 2008; Dediu & Ladd, 2007). Still in its beginnings and complex to investigate, it seems that second language acquisition (and hence what we can observe as adult second language learning aptitude) is subserved to a higher degree by heredity and hereditable factors than first language acquisition (Dale et al., 2010). This interesting result could be due to the fact that massive exposure time and experience with native languages overrides genetic influences and "levels them out", influences and differences which would potentially have been there in the first place as well. Not only genetic influences on second language acquisition or aptitude have recently been accumulated, but a steadily increasing body of research seems to emerge investigating the neural substrates of individual differences in expertise and success of foreign and second language learning, mostly in adult language learners. Brain function mostly via fMRI (functional magnetic resonance imaging) (Golestani & Zatorre, 2004; Hu et al., 2013; Kepinska et al., 2016; Reiterer et al., 2011) or EEG (electroencephalography) (Dogil & Reiterer, 2009), brain network states – so-called connectivity patterns or even "resting-state" patterns (fMRI, EEG) (Chai et al., 2016; Kepinska et al., 2017a, b; Prat, Yamasaki, Kluender, & Stocco, 2016) – as well as brain anatomy and brain structure (via MRI or DTI, diffusion tensor imaging) (Golestani & Pallier, 2007; Reiterer et al., 2011; Vaquero, Rodriguez-Fornells, & Reiterer, 2017; Turker et al., this volume) differences linked to individual differences are more and more investigated, and potential brain markers or "predictors" of language aptitude or language learning abilities in general are discerned and described. This recent cognitive neuroscience upstream in individual differences research concerning language abilities as one of the important cognitive abilities is mirrored in an increased interest within the field of SLA proper as well (Biedron, 2015; Darcy, Mora, & Daidone, 2016; Granena & Long, 2013; Safronova & Mora, 2012; Wen, Biedron, & Skehan, 2017).

In our own research focussing more on phonetic and speech imitation aptitude, apart from brain markers, we found markers in other psycho-cognitive domains. Higher speech imitation aptitude in adults and children was accompanied first and foremost by higher singing abilities but also higher general musicality and auditory working memory (Christiner, this volume; Christiner & Reiterer, 2015, 2013; Nardo and Reiterer, 2009), increased openness to new experience and empathy as personality markers (Hu et al., 2013; Rota & Reiterer, 2009) and differed between the sexes – with males showing elevated speech imitation skills and females showing superiority in grammar and vocabulary learning aptitude (Wucherer & Reiterer, 2016). As a phonetic marker of pronunciation aptitude for English as a second language, we could repeatedly isolate the initial schwa sound, mostly in content words, as a good predictor of overall pronunciation ability in L2 (Ghafoorian, this volume); we found minor markers in knowledge of multiple L1 dialects and increased speech imitation ability in L2; finally, we found very low to no correlations between L2 phonetic imitation aptitude and general nonverbal IQ, reading speed and executive functions. Last but not least, we always found all language aptitude subcomponents (e.g. phonetic, grammatical, lexical, pragmatic) to be normally distributed.

However, because the phenomenon of language aptitude is highly complex, influenced by many domains and factors (social, genetic, neuroscientific, psychological, cognitive and the like), we strived to explore it further by means of the manifold research projects comprised in this volume, by looking at many different factors to hopefully shed more light onto this complex phenomenon, which was a forgotten research field during the last decades before 2000, but now no longer is.

Thus, we are enthusiastic and hopeful that the field of aptitude research for foreign and second language learning will develop tremendously in the future years again.

Vienna, Wien, Austria Susanne Maria Reiterer

Bibliography

Biedron, A. (2015). Neurology of foreign language aptitude. *Studies in Second Language Learning and Teaching, 5*(1), 13–40. https://doi.org/10.14746/ssllt.2015.5.1.2

Chai, X. J., Berken, J. A., Barbeau, E., Soles, J., Callahan, M., Chen, J. K., & Klein, D. (2016). Intrinsic functional connectivity in the adult brain and success in second-language learning. *Journal of Neuroscience, 36*(3), 755–761. https://doi.org/10.1523/JNEUROSCI.2234-15.2016

Christiner, M., & Reiterer, S. (2013). Song and speech: examining the link between singing talent and speech imitation ability. *Frontiers in Psychology, 4*, 874. https://doi.org/10.3389/fpsyg.2013.00874

Christiner, M., & Reiterer, S. (2015). A mozart is not a pavarotti: singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience, 9*, 482. https://doi.org/10.3389/fnhum.2015.00482

Christiner, M., & Reiterer, S. (2016). Music, song and speech. A closer look at the interfaces between musicality, singing and individual differences in phonetic language aptitude. In G. Granena, D. O. Jackson & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 131–156). Amsterdam: John Benjamins.

Dale, P. S., Harlaar, N., Haworth, C., & Plomin, R. (2010). Two by two: A twin study of second language acquisition. *Psychological Science, 21*(5), 635–640. https://doi.org/10.1177/0956797610368060

Darcy, I., Mora, J. C., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning, 16*(4), 741–773. https://doi.org/10.1111/lang.12161

Dediu, D. (2008). The role of genetic biases in shaping the correlations between languages and genes. *Journal of Theoretical Biology, 254*(2), 400–407. https://doi.org/10.1016/j.jtbi.2008.05.028

Dediu, D., & Ladd, R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *PNAS, 104*, 10944–10949. https://doi.org/10.1073/pnas.0610848104

Dogil, G., & Reiterer, S. (2009). *Language talent and brain activity*. Berlin, Germany/New York: Mouton De Gruyter.

Golestani, N., & Pallier, C. (2007). Anatomical correlates of foreign speech sound production. *Cerebral Cortex, 17*(4), 929–934. https://doi.org/10.1093/cercor/bhl003

Golestani, N., & Zatorre R. J. (2004). Learning new sounds of speech: reallocation of neural substrates. *NeuroImage, 21*(2), 494–506. https://doi.org/10.1016/j.neuroimage.2003.09.071

Granena, G., & Long, M. (Eds.) (2013). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: John Benjamins.

Hayiou, T., Dale, P. S., & Plomin, R. (2012). The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Developmental Science, 15*(2), 233–249. https://doi.org/10.1111/j.1467-7687.2011.01119.x

Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., & Reiterer, S. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: behavioural predictors and neural substrates. *Brain & Language, 127*, 366–376. https://doi.org/10.1016/j.bandl.2012.11.006

Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. O. (2017a). On neural correlates of individual differences in novel grammar learning: An fMRI study. *Neuropsychologia, 8*(2), 156–168. https://doi.org/10.1016/j.neuropsychologia.2016.06.014

Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. O. (2017b). Whole-brain functional connectivity during acquisition of novel grammar: Distinct functional networks depend on language learning abilities. *Behav Brain Res, 320*, 333–346. https://doi.org/10.1016/j.bbr.2016.12.015

Marusakova, M. (2015). *Pragmatics and empathy in second language aptitude*. Master Thesis, University of Vienna.

Nardo, D., & Reiterer, S. (2009). Musicality and phonetic language aptitude. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 213–256). Berlin, Germany: Mouton de Gruyter.

Neufeld, G. (1979). Towards a theory of language learning ability. *Language Learning, 29*(2), 227–241.

Prat, C. S., Yamasaki, B. L., Kluender, R. A, & Stocco, A. (2016). Resting-state qEEG predicts rate of second language learning in adults. *Brain and Language, 157–158*, 44–50. https://doi.org/10.1016/j.bandl.2016.04.007

Reiterer, S., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., Winkler, S., & Ackermann, H. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: functional neuro-imaging and brain morphology. *Frontiers in Psychology, 2*(271):1–12. https://doi.org/10.3389/fpsyg.2011.00271

Reiterer, S., Hu, X., Sumathi, T. A., & Singh N. C. (2013). Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in Psychology, 4*, 782. https://doi.org/10.3389/fpsyg.2013.00782

Rota, G. & Reiterer, S. (2009). Cognitive aspects of pronunciation talent: how empathy, mental flexibility, working memory and intelligence interact with phonetic talent. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 67–96). Berlin, Germany: Mouton de Gruyter.

Safronova, E., & Mora, J. C. (2012). Acoustic and phonological memory in L2 vowel perception. *Proceedings of 22nd EUROSLA*, 384–390.

Vaquero, L., Rodriguez-Fornells, A., & Reiterer, S. M. (2017). The left, the better: White-matter brain integrity predicts foreign language imitation ability. *Cerebral Cortex, 27*(8), 3906–3917. https://doi.org/10.1093/cercor/bhw199

Wen, Z., Biedron, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching* (CUP), *50*(1), 1–31. https://doi.org/10.1017/S0261444816000276

Wucherer, B., & Reiterer, S. M. (2016). Language is a girlie thing, isn't it? A psycholinguistic exploration of the L2 gender gap. *International Journal of Bilingual Education and Bilingualism, 19*, 1–17. https://doi.org/10.1080/13670050.2016.1142499

Acknowledgement

We want to start by thanking the editing team without whom this book project would not have been possible. In particular, we are grateful to Victoria Ameringer, head of the team for editing, proofreading and general assistance, who contributed tremendously to the implementation of this project by organising and managing both the contributions inside this book and the correspondence with authors. Our special thanks and appreciations also go to the members of the editing and research team, Daniel Leisser, Luke Green and Sabrina Turker, who invested their time and expertise to substantially improve the quality of the manuscript. Special thanks go to Maria Handelberger for proofreading. Moreover, we praise the authors who shared their pearls of wisdom with us in order to create a book which provides a source of knowledge to students, researchers and teachers alike. We want to express our gratitude to Springer International Publishing AG for recognising the potential of our idea and for providing guidance throughout the whole process. Finally, we thank everyone who supported this project and "offered their help" to make this publication possible.

Contents

Contributors

Victoria Ameringer Department of Linguistics & Department of English and American Studies, University of Vienna, Vienna, Austria

Markus Christiner Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

Luke Green Department of English and American Studies, University of Vienna, Vienna, Austria

Cornelia Habl Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

Hannah Hackl Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

Sofia Hörder Department of English and American Studies, University of Vienna, Vienna, Austria

Klara Kager Department of Linguistics, University of Vienna, Vienna, Austria

Hyun Jung Kim Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

Katharina Krumpeck Department of Linguistics, University of Vienna, Vienna, Austria

Astrid Elisabeth Lehner Department of English and American Studies, University of Vienna, Vienna, Austria

Daniel Leisser Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

Zhaleh Ghafoorian Maddah Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Wien, Austria

Daniel Malzer Department of English and American Studies, University of Vienna, Vienna, Austria

Jakob Poschner Department of Linguistics & Department of English and American Studies, University of Vienna, Vienna, Austria

Susanne M. Reiterer Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Wien, Austria

Karin Richter Department of English and American Studies, University of Vienna, Vienna, Austria

Nejra Rizvanović Department of Philosophy, MEi:CogSci Programme, University of Vienna, Vienna, Austria

Stefanie Rüdegger Department of Linguistics, University of Vienna, Vienna, Austria

Niloufar Saraei Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

Peter Schneider Department of Neurology, Division of Neuroradiology and Section of Biomagnetism, University of Heidelberg, Heidelberg, Germany

Annemarie Seither-Preisler Centre for Systematic Musicology, Karl-Franzens University Graz & BioTechMed, Graz, Austria

Sabrina Turker Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

Introduction: Towards an Interdisciplinary Understanding of Language Aptitude

Victoria Ameringer, Luke Green, Daniel Leisser, and Sabrina Turker

1 Introduction to the Volume

This book represents a collection of quantitative studies on language aptitude in the context of psychology, the language sciences, and cognitive neuroscience. Throughout this book the authors bring together interdisciplinary approaches to language aptitude and its constituents in order to both inspire researchers and contribute to the testing of a new language aptitude test. This is because the authors are part of a teaching and research initiative by the editor, Susanne Maria Reiterer, embedded into the unit of language learning and teaching research (Sprachlehr- und –Lernforschung, SLLF), an interdisciplinary sub-unit affiliated with the linguistics department and the centre for teacher education at the University of Vienna. This research initiative uses a heuristic, exploratory bottom-up approach, as the ideas of the mixed group of young and senior investigators regarding new approaches on language aptitude research form the basis of this book. Some of the authors are part of the "aptitude and multilinguality" group which is currently initiating and exploring a new multilingual aptitude (MULT/AP) test that was developed by Markus Christiner and Susanne Reiterer at the University of Vienna (Christiner & Reiterer,

V. Ameringer (✉)
Department of Linguistics & Department of English and American Studies, University of Vienna, Vienna, Austria
e-mail: s1679742@ed.ac.uk

L. Green
Department of English and American Studies, University of Vienna, Vienna, Austria
e-mail: luke.green@univie.ac.at

D. Leisser · S. Turker
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria
e-mail: daniel.leisser@univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_1

2017). The MULT/AP test measures phonetic aptitude and aims at identifying factors influencing it, such as cognitive abilities, musicality, gender, and age. Hence, this book is a product of the joint effort of language aptitude researchers who aim at exploring language aptitude's multiple components, inspire Europe-wide research, and support the development of a novel language aptitude test.

Since language aptitude involves a myriad of constituents, we aimed at giving structure to the collection of articles in this book by organising them according to factors they have in common. The first factor that is explored in this book is language aptitude in the context of memory research. While Ameringer explores this topic with a broad approach that includes different memory functions and their impact on language aptitude, Hackl and Kim explore aptitude of vocabulary acquisition and the influence of WM in L2 vocabulary acquisition respectively. Secondly, the effect of psychological factors on language aptitude is investigated: Leisser examines the role of self-efficacy as a possible component of language aptitude in the acquisition of the British vowel /æ/, and Riznanović explores the influence of both personality and motivation. The book continues by explaining language aptitude in relation to neuroscience and musicality. The role of auditory cortex morphology in language aptitude is investigated by Turker et al. Kager explores language aptitude and its relation to hemispheric dominance, handedness, IQ, and game preferences. Christiner examines the relationship between music and language aptitude in pre-school children, Saraei focuses on the impact of speaking a tone language on music aptitude, and Malzer investigates musicality and grammar aptitude. Finally, Habl explores possible links between language aptitude and gender. This volume also looks at factors like socio-environmental influences, bilingualism and language attrition. While vocabulary acquisition strategies are central to Poschner's section, Krumpeck compares language aptitudes and language attitudes of monolingual and bilingual Burgenland Croats, Hörder scrutinizes the correlation of early multilingualism and language aptitude, and Lehner investigates the role of language aptitude in second language attrition. The final section in this chapter, written by Rüdegger, explores language aptitude as a concept to determine language talent. In the last chapter, pronunciation is the focus with the sections by Richter and Ghafoorian. The former investigates factors affecting the pronunciation abilities of adult learners of English with a longitudinal study, whereas the latter explores individual differences and aptitude in L2 phonology for the first time in Persian-speaking learners of English.

This structure reflects many of the similarities shared by the individual studies in this volume in addition to the common focus on language aptitude. For instance, Ameringer, Hackl, and Kim all investigate the impact of memory on language aptitude, while both Leisser and Riznanović take into account psychological influences, such as self-efficacy, motivation and personality. Krumpeck and Hörder investigate the impact of bilingualism or multilingualism on language aptitude. Christiner, Turker, Saraei, and Richter examine the correlation language aptitude shares with pronunciation and musicality. Finally, teaching methods and links to educational situations are identified by Poschner, Leisser, and Riznanović.

Due to the great variety of disciplines included in this volume, from the social sciences to the neurosciences, it is unsurprising that discrepancies regarding the fundamental 'nature or nurture' debate arise. For instance, while the majority of studies considers both 'nature' and 'nurture' to be crucial factors influencing language aptitude, those which concentrate on genetics and neuroscience rather advocate a 'nature' approach to variances in language aptitude, while those chapters focusing on the social environment of language learners tend to support a stance of 'nurture'. Furthermore, the studies in this book are distinct in their primary purposes as each chapter suggests implications for different fields. While some of the studies are conducted in order to enhance pedagogical techniques, other researchers predominantly aim at finding novel results or at reconstructing existing hypotheses to confirm previous findings. Notably, many studies are able to provide new, significant findings which both aid teachers and disclose previously undiscovered issues.

As evident in the structure, this book is unique in that it explores language aptitude from multiple perspectives, thereby allowing the reader to obtain a view of the subject matter which is both broad and confined simultaneously. Although the articles are distinct with regard to their individual foci and research questions, they share clear commonalities, both in terms of their relevance to the broader field of language aptitude, as well as their aim towards either developing pedagogical techniques or producing new results. The introduction to this volume will continue by outlining perspectives for future research within the field of language aptitude, before providing a brief overview of foreign language aptitude research.

2 Implications and Future Research

Since language aptitude as a research field is rather young (Neufeld, 1979), a number of problems arose for the authors while devising their studies. Firstly, some topics, such as language attrition and its connections with language aptitude, have been scarcely investigated, which made it a challenging endeavour to compare the existing literature. Secondly, due to a lack of research in some areas of language aptitude, the reasons for the occurrence of some significant results were subject to speculation. Future research should therefore replicate the studies in this book, especially those yielding novel findings, in order to confirm the results and potentially draw similar conclusions.

Additionally, the studies in this book could be enhanced by refining the selection of study participants, as researchers were sometimes limited in their power to choose the best possible candidates. Future replication projects would certainly benefit from the investigation of a higher number of individuals to increase both the validity and reliability of findings. Furthermore, the 'nature and nurture' approach supported by most authors in this book seems to justify the call for a larger number of longitudinal studies involving to test and re-test individuals' language aptitude over a sufficient amount of time.

Despite the problems encountered by the authors, this volume produces findings which provide implications for teachers, students, and researchers alike. For instance, Poschner finds that vocabulary acquisition strategies do not differ much between high and low aptitude learners, which implies that the effectiveness of such learning strategies with respect to increasing cognitive abilities (especially towards the higher end of the aptitude distribution spectrum) can be questioned. Similarly beneficial are Leisser's results, which demonstrate that self-efficacy is a crucial element of language learning and that it significantly correlates with phonetic aptitude. Students should profit from a positive mind set when learning a foreign language, since their beliefs regarding their own capabilities of producing a sound in a native-like manner can influence their performance.

What is more, the following studies might offer implications to researchers, as they either challenge or support current hypotheses, and introduce novel, empirical findings: Ameringer discovers the beneficial effects of tertiary education and demonstrates that the complex thinking this high level education requires enhances memory functions and language aptitude. Lehner adds to the scarce body of literature investigating the relationship between language attrition and language aptitude, and finds that a high language aptitude compensates for infrequent L2 input. According to Lehner, this is because only low aptitude learners are affected by the duration of L2 learning, the length and extent of attrition time, and age of onset of L2 learning. Although Riznanović cannot support her two main hypotheses regarding the positive influence of high empathy and conscientiousness on language aptitude, she does find that intrinsic motivation and a phlegmatic temperament significantly benefit language acquisition. Leisser connects the notions of aptitude and attitude, suggesting that high phonetic coding ability may be linked to a higher success rate in natural selection. He argues that longitudinal studies on the linguistic socialisation of individuals starting before second language onset may provide interesting insights. Such studies may enable researchers to collect data on both psychological variables and language aptitude at the same time. Finally, the findings in Christiner's study may be investigated according to gender-specific differences with regard to children's musical abilities and their speech imitation talent.

This book highlights the importance of scrutinizing language aptitude by revealing that its interdisciplinarity enables researchers to generate findings which do not only provide significant insights into individual differences in language learning but also into pedagogy, neuroscience, genetics, linguistics, and psychology, all of which can contribute to this small subfield of SLA. In a next step, language aptitude research should also strike roots in legal and medical contexts so as to provide assistance in the process of designing and evaluating language-related test constructs applied in admission tests such as the Law National Aptitude Test (LNAT) or the BioMedical Admissions Test (BMAT), e.g. verbal reasoning abilities, reading comprehension and pragmatic decoding of information.

As can be seen in this book's articles that explore language aptitude from a multitude of perspectives, language aptitude per se is a highly complex, multi-faceted construct. This volume provides its readers with an exciting journey into the field of aptitude research and helps them get a better understanding of the complexity of the concept. Before starting this journey, however, we advise our readers to go carefully through the following up-to-date introduction dealing with the theoretical construct of language aptitude. It is a thorough review of recent literature attempting to get to the core of foreign language aptitude (FLA) research by commenting on its nature, complexity and its components.

3 Getting to the Core of Language Aptitude

It may not come as a surprise that finding a uniform definition of *language aptitude* is just as challenging as finding one simple, compact definition of concepts like *intelligence* or *self-concept*. Thanks to the renewed interest it has received in past years, the concept of FLA has undergone considerable change. This being said, there is not one single sentence that could capture the whole concept of language aptitude successfully and, adding to the difficulty, there is still considerable debate regarding its stability, innateness, complexity and particularly its components (Li, 2015, 2016). It might sound like a challenging, if not even impossible, undertaking to summarize the most recent developments on FLA on but a few pages, yet there certainly is a need for a clear and precise theoretical investigation before reading the studies presented in this book.

3.1 A Brief History of FLA

Scientific interest in language aptitude as an aspect of language acquisition that needs to be investigated did not arise before the beginning of the twentieth century. In Europe, interest arose much earlier, but the first steps of language aptitude research were taken in the United States, where almost all testing batteries were developed. There, colleges and universities first started to show an interest in the language skills of their students in the 1920s, which can be seen as the birth of objective testing. It was not before 1960 that the government started investing in this trend, albeit mainly for political and social reasons. To sum up, proficiency or aptitude was first only used to assess linguistic competence on an academic level for mostly practical purposes and the second wave of language aptitude research in the 1950s and 1960s saw the birth of the most widespread aptitude batteries. Mainly based on the assumptions made by John B. Carroll (1958, 1962, 1973, 1990), the tests designed in the early years of language aptitude measured whether individuals were fast or slow learners of foreign languages and were not based on any theoretical framework (intelligence tests containing linguistic elements are left out at this point) (Spolsky, 1995).

FLA is a concept that dates back to the beginning of the second half of the twentieth century and for almost half a century, it was defined by the rate of acquisition at which an unknown language was learned (Carroll, 1958, 1962, 1973, 1990; Stansfield & Reed, 2004). To put it in simpler terms, those who learnt/acquired a language with a certain speed, i.e. very quickly, were considered to have a certain aptitude for learning foreign languages. This assumption was first introduced by Carroll, who is often also referred to as the father of FLA research (Spolsky, 1995). According to Spolsky, he was first and foremost one of the founders of the discipline of psycholinguistics and coined the American psychological and linguistic world from 1950 onwards. Being the first graduate student of B.F. Skinner and later on a Harvard professor, he organized the first seminars in the course of which the psychology of language was claimed a separate, interdisciplinary field. Carroll became most famous for his well-known Modern Language Aptitude Test (henceforth MLAT), which he designed together with his colleague Stanley Sapon (1959). Since then, the theoretical construct of language aptitude has been questioned and modified by numerous researchers, yet the core characteristics introduced by Carroll, such as the fact that FLA consists of various components and depends on a number of external and personal factors, are still upheld today. One of the main problems of FLA research after the 1970s was that it was partly perceived as anti-egalitarian, and testing an individual's ability to learn something had a negative connotation, being tinged with forbiddance (Skehan, 2002). It was not until recently that researchers have picked up the topic and have revived the debate as to what FLA actually is and what it consists of.

Looking back, it can be said that it was not before the beginning of the twenty-first century that research focusing on aptitude and individual differences generally, but also on language aptitude and its various subcomponents, gained increasing interest in numerous fields such as education, didactics, cognitive neuroscience and cognitive psychology. The research area has truly experienced a boom and has gained considerable momentum. But what is it that has become so popular, and do we even know what we are exploring when conducting studies on FLA?

3.2 What Is Foreign Language Aptitude?

Foreign language aptitude is a term that subsumes a number of concepts and is often used interchangeably with other terms, such as talent, giftedness language learning ability or even sometimes with language learning expertise. Although it is often still difficult to know where to draw the line and differentiate the variety of terms, researchers have at least suggested a differentiation between talent and aptitude according to which aptitude designates the innate property that develops into a certain skill, which is then termed *talent* (Gagné, 1995, 2005; Stern & Neubauer, 2013). According to Gagné (2005), who developed the so-termed Differentiated Model for Giftedness and Talent, *giftedness* and *talent* are two terms that have to be clearly distinguished. The term *giftedness* can be synonymously used for aptitude and refers to an undeveloped, biologically inherited predisposition for acquiring a certain skill. In other words, giftedness is what people are born with, and talent is what an individual develops out of aptitude and the proficiency they achieve. This shows that talent is a skill that develops over time and is not already stable or fixed at birth. To be more precise, high aptitudes become well-trained skills (expertise) that are systematically developed (Gagné, 2005; Seither-Preisler, Parncutt, & Schneider, 2014).

First of all, it is important to note that even if language aptitude is not used in combination with "foreign", i.e., foreign language aptitude, it never applies to first language acquisition or bi/tri/multilingual language learning (multilingualism). The term is usually only applied to the acquisition of novel languages. Carroll (1962, 1990) understands language aptitude to be a state of readiness of individuals which provides them with a certain capacity and facility for learning foreign languages if motivation and opportunity are provided. The aforementioned definition proposed by Carroll describes FLA as an innate trait that is immune to training and is stable over time, an aspect that has faced quite some criticism in the past years (Dörnyei, 2006; Robinson, 2001, 2002, 2005, 2012; Singleton, 2017; Skehan, 2002; Wen, Skehan, & Biedroń, 2017). Nowadays, the general held view is that language aptitude is a more dynamic, multi-faceted conglomerate of various cognitive skills (Biedroń, 2015; Dörnyei, 2006) that can under certain circumstances be altered through practice (Kormos, 2013; Singleton, 2017).

Robinson (2005) tries to integrate this shift by describing FLA as "[s]econd language (L2) learning aptitude […] characterized as strengths individual learners have—relative to their population—in the cognitive abilities information processing draws on during L2 learning and performance in various contexts and at different stages" (p. 46). Robinson (2001, 2002, 2005, 2012) acknowledges the importance of FLA as a strength an individual may possess compared to peers but also points towards the significance of taking other factors into account. Learning a foreign language requires high cognitive abilities, e.g. working memory, metalinguistic awareness (Jessner, 2006, 2014; Singleton, 2017), and never takes place out of a context. Therefore, the context of learning, as well as the stage at which an individual starts acquiring a foreign language certainly play a role.

More recently, the claim that language aptitude may change and be modified by other factors, has gained considerable attention in FLA research. The debate as to whether FLA is (partly) innate or just one of the many factors influencing successful foreign language acquisition is still troubling researchers from various disciplines. As Kormos (2013) puts it,

> [a]lthough language-learning aptitude might seem to be a relatively stable individual characteristic […] there seems to be some converging evidence that certain components of aptitude […] might improve in the course of language learning. (p. 145f)

The construct of language aptitude has also served as a topic of debate due to the rising importance that working memory has gained in recent years (Biedroń, 2015; Biedroń & Pawlak, 2016a, 2016b; Wen, 2012, 2016; Wen et al., 2017). Owing to the strong correlation found between these two skills, some researchers have even proposed that working memory capacity could be seen as equivalent to language aptitude (Miyake & Friedman, 1998; Sawyer & Ranta, 2001; Wen, 2016; Wen & Skehan, 2011; Wen et al., 2017). Studies investigating the two have confirmed the impact of working memory on numerous linguistic abilities, e.g. faster and often more successful first and foreign language acquisition (Ellis & Sinclair, 1996;

Kormos & Sáfár, 2008; Linck et al., 2013; Miyake & Friedman, 1998; Sáfár & Kormos, 2008). Better working memory skills thus seem to lead to more successful foreign language learning (Biedroń, 2012; Van den Noort, Bosch, & Hugdahl, 2006). It is noteworthy, however, that there are striking differences between specific working memory components, their measurability and to which extent they relate to the known components of foreign language aptitude (Baddeley, 2003a, 2003b, 2017; Jacquemot & Scott, 2006). In addition, other studies have questioned the suggested impact working memory has been claimed to have on language aptitude (Winke, 2013).

3.3 The Components of Language Aptitude

The four major components of language aptitude claimed by Carroll (1958, 1962, 1973) are (1) Phonetic Coding Ability, (2) Grammatical Sensitivity, (3) Inductive Language Learning Ability and (4) Rote Learning Ability. Although these components date back to the 1960s, they are to a great extent still upheld today. Theoretical advancements have been proposed by a number of researchers, such as Peter Robinson and Peter Skehan. Biedroń (2015) highlights that the heterogeneity of the construct is certainly one of the main obstacles that researchers encounter when trying to investigate FLA. Researchers in the past decades have followed Carroll's classic model of four abilities subsumed under FLA but many have added or modified them.

Robinson (2001, 2002, 2005, 2012) postulates a theory of so-called aptitude complexes, which include neglected factors such as the interaction between different aptitude components and the implications that might be drawn for foreign language learning circumstances. He focuses on the pedagogical side, denies a hierarchical structure of aptitude and states that "sets of cognitive abilities, or 'aptitude complexes' are differentially related to language learning under different psycholinguistic processing conditions" (Robinson, 2001, p. 369). For an in-depth account on the distinction of abilities (e.g. primary or core abilities) and his theoretical approach, please refer to Robinson (2007). Skehan (1986, 1991, 2002) also thoroughly investigates the development of FLA in his works and distinguishes between three rather than four categories. He refers to grammatical sensitivity and inductive language ability as one form of language analytic ability, whereas the first is passive, the latter active. He applies these three components to three models, which for him are instantly connected, namely auditory processing (phonetic coding ability), language processing (language analytic ability) and memory (associative memory). More generally he speaks about these modularity influences in second language learning processes but given that language aptitude is supposed to facilitate this process, it directly implies that the three processes must be the reason for high proficiency, for instance.

Skehan has succeeded in merging inductive language learning ability and grammatical sensitivity into one category, now commonly termed *language analytic* *ability* (Abrahamsson & Hyltenstam, 2008; Biedroń, 2015; Biedroń & Pawlak, 2016a, 2016b; Kepinska, de Rover, Caspers, & Schiller, 2017; Kocić, 2010; Robinson, 2001, 2002, 2012; Wen et al., 2017). Claims have recently also been made that the components of language aptitude might be relevant for different stages and contexts of learning (Abrahamsson & Hyltenstam, 2008; Artieda & Muñoz, 2016; Hu et al., 2013), but very little research has been conducted supporting this view so far.

In the past years, all external and internal influences that may impact foreign language learning have been emphasized, in particular in pedagogical domains. Although providing a clear picture of all recent developments would go far beyond the scope of this introductory chapter, the words of Biedroń and Pawlak (2016a) provide a fitting summary:

> A gifted FL learner is a person who, owing to his/her exceptional inborn gift for learning languages, especially capacious verbal working memory, as well as expertise in learning foreign languages, is able to learn any foreign language to a near-native level of competence, given proper motivation, time and conducive environment. (p. 155)

This quote already gives an impressive example of the complexity of the concept and what we can say for sure is that language aptitude is not a uniform concept and it certainly cannot be measured using one single test. Even more importantly, it cannot be expressed through one single score, i.e. there is no black and white in language aptitude (testing). Different factors have an influence on the development of language aptitude in an individual and there is common agreement that language aptitude consists of a number of rather distinguished skills. Language aptitude is only one of the many factors accounting for the individual differences found in SLA research, but it certainly plays a vital role.

3.4 Aptitude Testing

The best known and most widely used language aptitude test is the MLAT, which has also been used in some research projects presented in this book. Together with his colleague Sapon, Carroll (1959) chose a purely empirical approach and selected forty different cognitive and psychological tests. The tests that actually matched the language proficiency scores were then chosen to make up the so-called Modern Language Aptitude Test – an approach that had been frequently applied in intelligence testing.

Robinson (2002) criticizes that the MLAT was funded and developed with the aim of predicting the rate of acquisition in as simple and effortless a testing as possible, i.e. paper-and-pencil methods in one sitting with one single, definite score. The major purpose of this process was the adequate categorization of individuals in different programs based on their potential, defined by the score achieved. Likewise, Pimsleur's Language Aptitude Battery (PLAB) (Pimsleur, 1966) was mainly designed to meet these purposes, but the main difference was that it was administered to a younger population. In sum, the MLAT, the PLAB, the VORD and the DLAB (Defense Language Aptitude Battery by Petersen & Al-Haik, 1976) are all of similar nature and were all developed for a similar purpose, namely to differentiate between students who had very high proficiency and those with very low proficiency.

At the beginning of the twenty-first century, Grigorenko, Sternberg, and Ehrmann (2000) presented a new theory of foreign language aptitude and a testing battery measuring their theoretical construct, the Cognitive Ability for Novelty in Acquisition of Language (Foreign) (CANAL-F). According to the model of the three authors, there are four aspects essential for language aptitude, namely (1) Knowledge Acquisition Processes, (2) Levels of Processing, (3) Modes of Input, and (4) Memory Processes (Rysiewicz, 2008) which operate through different processes. What is striking about the CANAL-F is that it is dynamic given that the individual learns a new (artificial) language while doing the test. The language is called Ursulu and resembles natural languages, but does not directly correspond to any one existing language.

A rather recent and very useful language aptitude test is the LLAMA (Meara, 2005). This computer-based testing battery is a revised version of the Swansea LAT and, based on the components introduced by Carroll (1959, 1962), measures language aptitude through four different tests (Granena, 2013). It has certainly gained popularity and, as Granena points out, only the LLAMA test does not suffer from any limitation or restriction, e.g. being difficult to get, being available only in pencil-and-paper format, or only being used for military purposes. Even though it is not extensively standardized like the MLAT, it has been used by numerous research groups and is available for free online. As Granena describes, and as can be read in the LLAMA manual (Meara, 2005; Rogers, Barnett-Legh, Curry & Davey, 2015), the LLAMA is based on the MLAT and is language-independent (based on material from an indigenous Central American language), which is a major advantage over other tests such as the MLAT, which is only available in a very limited number of languages. Language independence is a feature of great importance that has numerous advantages, e.g. facilitation of test administration, no influences from native language or other languages spoken by participants (Granena, 2013; Rogers et al., 2015).

The four subtests of the LLAMA, following the scheme of the MLAT, are:

1. LLAMA B: a test of vocabulary learning
2. LLAMA D: a sound recognition task
3. LLAMA E: sound-symbol association
4. LLAMA F: a test measuring grammar inferencing

The LLAMA is also the test that will appear most frequently in this volume due to the various advantages just mentioned. One last testing battery that deserves to be mentioned is the so-called Hi-LAB, which was invented for military purposes with funding from the US government. Very few papers have been published so far (Linck et al., 2013) explaining the relevance and reliability of the test, but little can be said about it, since it is not publicly available and for scientific purposes (following rumours) only with extremely limited access. Clarification regarding this issue is still awaited.

3.5 *Caveats of Aptitude Testing*

At this point it should be mentioned that despite the quite impressive theoretical and practical advancements in many areas of psycholinguistic research in the past decades, language aptitude research has seen very little progress with respect to the development of adequate testing batteries. The main reasons for this moderate progress are probably the time-consuming nature of such an undertaking, a lack of access to established testing batteries for researchers and the lack of a clear and precise theoretical underpinning of the construct of language aptitude. Fortunately, the topic has gained increasing attention in the past years and researchers have partly succeeded in updating the theoretical construct in accordance with most recent results of psycholinguistic research (Christiner & Reiterer, 2017; Christiner, Rüdegger, & Reiterer, 2018; Reiterer, forthcoming). Still, researchers have not been successful at developing more reliable, powerful and up-to-date full batteries compared to the traditional tests, such as the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959).

The MLAT, the PLAB, the VORD (for details, see Parry & Child, 1990) and the Defense Language Aptitude Battery (DLAB; Petersen & Al-Haik, 1976) are all of similar nature and were all developed for a similar purpose, namely differentiating between students and individuals who had very high linguistic potential and those with very low aptitude (Robinson, 2002). Interestingly, these tests were all based on the MLAT and relied on the fact that what Carroll had described as measuring language aptitude was correct (Carroll, 1962, 1989, 1990). What we can be certain about is that the MLAT definitely measures a variety of components involved in language aptitude but it lacks a theoretical foundation. Additionally, it remains questionable in how far the aforementioned components are stable and present at every stage of language learning. Researchers from pedagogic backgrounds, e.g. Dörnyei, Robinson and Skehan, have argued for a dynamic conceptualization of aptitude and proposed that access to the cognitive abilities associated with language aptitude may be important for different stages of learning (Dörnyei, 2005; Robinson, 2002, 2012; Skehan, 2002; Yalçin & Spada, 2016) – an aspect definitely worth taking into account in future research on language aptitude.

A further problematic aspect of language aptitude testing is the availability of the given tests. Obtaining a version of a famous language aptitude battery is similar to finding the needle in the haystack. The US government prefers to keep its tests secret and unavailable, which makes them seem like a myth to language aptitude researchers in other parts of the world. The DLAB and the VORD are protected tests, which are only administered to United States government personnel (Robinson, 2002). Those researchers seeking to obtain a version of the MLAT, which is claimed to be available to a broader audience and open for research, will find it just as difficult as with the DLAB and VORD. As can be read online, "Unfortunately, due to the sensitive nature of the test, we only sell the MLAT to government agencies, missionary groups, and licensed clinical psychologists. We do NOT sell the test to individual researchers, teachers, or students" (Language Learning and Testing

Foundation, 2014). Still, most papers on language aptitude cite the MLAT and state that it is a commercially available test (Yalçin & Spada, 2016). With respect to the Hi-LAB, it was invented for military purposes with funding from the US government, and access to this battery is usually not provided (information on the test can be found in Linck et al., 2013). Very few papers have been published so far explaining the relevance, validity and reliability of the test but, requests to the originators remain usually unanswered, although requesting single copies for research purposes is encouraged on the webpage. Finally, regarding the CANAL-F, requesting a copy of it for use in research is as problematic as with the others. This pencil-and-paper test cannot be obtained due to the fact that there is no online version and no paper-and-pencil versions are available, traceable anymore. Apart from that, it was also designed for use in government contexts (Robinson, 2002).

4 Conclusion

As demonstrated in this introduction, the field of language aptitude is highly interdisciplinary and has evolved drastically over the past decades. Although language aptitude testing has its caveats, it is nonetheless a crucial component in language aptitude research which has enabled progress. Additionally, a newer language aptitude test, the LLAMA, embodies a positive example of language aptitude testing as it does not have the disadvantage of other tests: it is free, easily accessible, language non-specific, and user friendly (Yalçın & Spada, 2016).

Although the progress which has been made in language aptitude research is impressive, there are still a plethora of questions to answer before this concept can be fully understood. This book sheds light on the internal workings of language aptitude by exploring it from various perspectives and fields. As will be evident in the following articles, some studies raised issues which could not be fully resolved since language aptitude research is still in its infancy. Thus, the interpretation of some results could not always be based on scientific evidence but on speculation. Nonetheless, this volume contributes substantially to a better understanding of language aptitude and will hopefully inspire its readers to join the research in this field.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in second language acquisition, 30*(4), 481–509. https://doi.org/10.1017/S027226310808073X

Artieda, G., & Muñoz, C. (2016). The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learning and Individual Differences, 50*, 42–48. https://doi.org/10.1016/j.lindif.2016.06.023

Baddeley, A. (2003a). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2003b). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2017). Modularity, working memory and language acquisition. *Second Language Research, 33*(3). https://doi.org/10.1177/0267658317709852

Biedroń, A. (2012). Memory abilities in gifted foreign language learners. In *New perspectives on individual differences in language learning and teaching* (pp. 77–95). Berlin/Heidelberg, Germany: Springer.

Biedroń, A. (2015). Neurology of foreign language aptitude. *Studies in Second Language Learning and Teaching, 5*(1), 13–40. https://doi.org/10.14746/ssllt.2015.5.1.2

Biedroń, A., & Pawlak, M. (2016a). New conceptualizations of linguistic giftedness. *Language Teaching, 49*(2), 151–185. https://doi.org/10.1017/S0261444815000439

Biedroń, A., & Pawlak, M. (2016b). The interface between research on individual difference variables and teaching practice: The case of cognitive factors and personality. *Studies in Second Language Learning and Teaching, 6*(3), 395–422. https://doi.org/10.14746/ssllt.2016.6.3.3

Carroll, J. B. (1958). A factor analysis of two foreign language aptitude batteries. *The Journal of General Psychology, 59*(1), 3–19.

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87–136). Pittsburgh, PA: University of Pittsburgh Press.

Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *Linguistics, 11*(112), 5–14.

Carroll, J. B. (1989). The Carroll model: A 25-Year retrospective and prospective view. *Educational Researcher, 18*(1), 26–31 Retrieved from http://www.jstor.org/stable/1176007

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude: Then and now. In T. Perry & C. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11–29). Englewood Cliffs, NJ: Prentice Hall.

Carroll, J. B., & Sapon, S. (1959). *Modern language aptitude test*. Bethesda, MD: Second Language Testing.

Christiner, M., & Reiterer, S. M. (2017). MULT/AP Multilingual aptitude test. Vienna: Christiner questionnaires. Retrieved from http://www.multap.net

Christiner, M., Rüdegger, S., & Reiterer, S. M. (2018). Sing Chinese and tap Tagalog? Predicting individual differences in musical and phonetic aptitude using language families differing by sound typology. *International Journal of Multilingualism, 3*, 1–17.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*(1), 42–68. https://doi.org/10.1075/aila.19.05dor

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology, 49*(1), 234–250.

Gagné, F. (1995). From giftedness to talent: A developmental model and its impact on the language of the field. *Roeper Review, 18*(2), 103–111.

Gagné, F. (2005). From gifts to talents: The DMGT as a developmental model. In J. E. Davidson & R. J. Sternberg (Eds.), *Conceptions of giftedness* (pp. 98–119). Cambridge, UK: Cambridge University Press.

Granena, G. (2013). Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 105–130). Amsterdam: John Benjamins Publishing Company.

Grigornko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal, 84*(3), 390–405. https://doi.org/10.1111/0026-7902.00076

Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., & Reiterer, S. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: Behavioural predictors and neural substrates. *Brain & Language, 127*, 366–376.

Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences, 10*(11), 480–486. https://doi.org/10.1016/j.tics.2006.09.002

Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language*. Edinburgh, UK: Edinburgh University Press.

Jessner, U. (2014). On multilingual awareness or why the multilingual learner is a specific language learner. In *Essential topics in applied linguistics and multilingualism* (pp. 175–184). Cham: Springer.

Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. O. (2017). On neural correlates of individual differences in novel grammar learning: an fMRI study. *Neuropsychologia, 98*, 156–168. https://doi.org/10.1016/j.neuropsychologia.2016.06.014

Kocić, A. (2010). The effects of aptitude on language learning. *Komunikacija I Kultura Online: Godina I, Broj, 1*(1), 234–243.

Kormos, J. (2013). New conceptualizations of language aptitude in second language attainment. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude and ultimate attainment* (pp. 131–152). Amsterdam: John Benjamins Publishing.

Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*(2), 261–271. https://doi.org/10.1017/S1366728908003416

Language Learning and Testing Foundation. (2014). *Language Aptitude Tests*. Retrieved from http://lltf.net/.

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics, 36*(3). https://doi.org/10.1093/applin/amu054

Li, S. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition, 38*(4). https://doi.org/10.1017/S027226311500042X

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., et al. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning, 63*(3), 530–566. https://doi.org/10.1111/lang.12011

Meara, P. (2005). *LLAMA language aptitude tests: The manual*. Swansea, UK: Lognostics.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Neufeld, G. G. (1979). Towards a theory of language learning ability. *Language Learning, 29*(2), 227–241.

Parry, T. S., & Child, J. R. (1990). Preliminary investigation of the relationship between VORD, MLAT and language proficiency. *DOCUMENT RESUME, 36*.

Petersen, C. R., & Al-Haik, A. R. (1976). The development of the defense language aptitude battery (Dlab). *Educational and Psychological Measurement, 36*(2), 369–380.

Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery (form S)*. New York: Harcourt, Brace and World, Incorporated.

Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research, 17*(4), 368–392. https://doi.org/10.1177/026765830101700405

Robinson, P. (2002). Effects of individual differences in intelligence, aptitude, and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfield and Hernstad (1991). In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211–265). Amsterdam: John Benjamins Publishing.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73. https://doi.org/10.1017/S0267190505000036

Robinson, P. (2007). *Aptitudes, abilities, contexts, and practice. Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. New York: Cambridge University Press.

Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes, and aptitude test development. In *New perspectives on individual differences in language learning and teaching* (pp. 57–75). Berlin/Heidelberg, Germany: Springer.

Rogers, V., Barnett-Legh, T., Curry, C., & Davey, E. (2015). *Validating the LLAMA aptitude tests*. York, UK: EUROSLA 2015.

Rysiewicz, J. (2008). Measuring foreign language learning aptitude. Polish adaptation of the Modern Language Aptitude Test by Carroll and Sapon. *Poznań Studies in Contemporary Linguistics, 44*(4), 569–595. https://doi.org/10.2478/v10010-008-0027-6

Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *IRAL – International Review of Applied Linguistics in Language Teaching, 46*(2). https://doi.org/10.1515/IRAL.2008.005

Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). Cambridge, UK: Cambridge University Press.

Seither-Preisler, A., Parncutt, R., & Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. *Journal of Neuroscience, 34*, 10937–10949. https://doi.org/10.1523/JNEUROSCI.5315-13.2014

Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute? Studies in Second Language Learning and Teaching, 7(1), 89-103. doi:https://doi.org/10.14746/ssllt.2017.7.1.5

Skehan, P. (1986). *Where does language aptitude come from?* Washington, DC: ERIC Clearinghouse.

Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition, 13*(2), 275–298.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–95). Amsterdam/New York: John Benjamins.

Spolsky, B. (1995). Prognostication and language aptitude testing, 1925–62. *Language Testing, 12*(3), 321–340.

Stansfield, C. W., & Reed, D. J. (2004). The story behind the modern language aptitude test: An interview with John B. Carroll (1916–2003). *Language Assessment Quarterly: An International Journal, 1*(1), 43–56. https://doi.org/10.1207/s15434311laq0101_4

Stern, E., & Neubauer, A. (2013). *Lernen macht intelligent: Warum Begabung gefördert werden muss*. München, Germany: DVA.

Van den Noort, M. W., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11*(4), 289–296. https://doi.org/10.1027/1016-9040.11.4.289

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting a case for "working memory as language aptitude". *Ilha Do Desterro: A Journal of English Language, Literatures in English and Cultural Studies, 60*, 015–044.

Wen, Z. (2012). Foreign language aptitude. *ELT Journal, 66*(2), 233–235. https://doi.org/10.1093/elt/ccr06887

Wen, Z. (2016). *Working memory and second language learning: Towards an integrated approach*. Bristol, UK: Multilingual Matters.

Wen, Z., Skehan, P., & Biedroń, A. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching, 50*(1), 1–31. https://doi.org/10.1017/S0261444816000276

Winke, P. (2013). An investigation into second language aptitude for advanced Chinese language learning. *The Modern Language Journal, 97*(1), 109–130. https://doi.org/10.2307/23361741

Yalçın, Ş., & Spada, N. (2016). Language aptitude and grammatical difficulty. *Studies in Second Language Acquisition, 38*(02), 239–263. https://doi.org/10.1017/S0272263115000509

Part I
Language Aptitude and Memory

Cognitive Abilities: Different Memory Functions and Language Aptitude

Victoria Ameringer

Abstract Past research on language aptitude has intensely focused on individual differences (IDs) since these are regarded as crucial predictors for an individual's innate ability to acquire a foreign language. This paper investigates memory as an essential ID, as it was found to be a fundamental element by previous research, and presents novel, empirical evidence on the influence that both memory and education have on language aptitude. Specifically, the primary foci lie on both verbal working memory (WM) and declarative memory (DM) capacity, and on the effects that both the length and type of education have on these memory systems. It is hypothesised that a greater capacity of verbal WM and DM leads to a higher language aptitude, and that a longer and a higher education level increases the capacities of both memory systems which, in turn, also augments language aptitude. Research was conducted by testing a homogenous sample of 30 participants, which was split into equally sized groups differing in educational status (workers without an academic background and university students). The analyses revealed that DM capacity is a predictor for language aptitude in contrast to verbal WM, which was against expectations due to previous findings in the field. Furthermore, both a longer and a higher education level were found to mutually increase these memory capacities and language aptitude.

1 Introduction

Since humans' minds are distinctively constructed and hence operate uniquely, it is self-evident that individual differences (IDs) have an influence on the development of cognitive abilities, such as second language acquisition (SLA) (Dörnyei, 2006). One of these IDs is a component labelled *language aptitude*, which embodies a

V. Ameringer (✉)
Department of Linguistics & Department of English and American Studies, University of Vienna, Vienna, Austria
e-mail: s1679742@ed.ac.uk

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_2

factor determining an individual's innate ability to acquire a foreign language (Granena & Long, 2013). Although the literature is indecisive with regard to an exact definition, language aptitude can be construed as a complex of elements comprising various abilities which enable language acquisition (Carroll & Sapon, in Dörnyei 2006). Expressed differently, language aptitude is a factor encompassing numerous sub-components whose measurement reveals an individual's "general language learning abilities" (Jilka, in Dogil & Reiterer, 2009, p. 21). Due to John B. Carroll, who is the initiator of the term *language aptitude* and co-founder of the Modern Language Aptitude Test (MLAT), the assessment of a successful language learning outcome is possible (Carroll, 1971).

One of the sub-components constituting language aptitude, namely memory, stands in the focus of this paper and will be thoroughly examined in order to explain variations in individual SLA. As language acquisition is dependent on a learner's ability of creating associations, Dual Coding Theory (DCT), i.e. a model explaining associative memory, will be described in order to provide essential background information which constitutes the frame of this paper. In addition, memory as a cognitive ability will be divided into the constituents working memory (WM), whereupon a special focus lies on verbal WM, and declarative memory (DM). Moreover, the variable *education* will be considered in the context of both language aptitude and memory to examine the possible influences of an environmental factor on the cognitive components under consideration. Even though (verbal) WM is already considered a crucial constituent of language aptitude (Baddeley, Gathercole, & Papagno, 1998; Miyake & Friedman, 1998; Papagno & Vallar, 1995; Wen & Skehan, 2011), the literature is less prevalent regarding the connections that language aptitude has with both DM and education. The motivation behind this paper lies therefore in the contribution to a better understanding of the impact that both DM and education have on language aptitude, and also in the further verification of previous findings regarding verbal WM. Hence, my research question is formulated accordingly: Do verbal WM, DM, and education influence language aptitude, and is there a relationship between the cognitive factors and the environmental component? In order to examine these questions, I developed four hypotheses which will be described, and tested, in the proceeding sub-sections after relevant background knowledge is provided.

1.1 Dual Coding Theory and Language Aptitude

Dual Coding Theory (DCT) was first introduced in 1971 by Allen Paivio, who was interested in the mental processes which underlie human cognition. His theory is generally concerned with "symbolic representational systems" (Paivio, 1986) which are capable of structuring and associating information from the environment. The core idea of DCT is that this environmental information can be encoded into memory in the shape of two differing memory codes, namely as verbal or non-verbal information (Reed, 2010, p. 126). According to Paivio (1986, p. 53), non-verbal

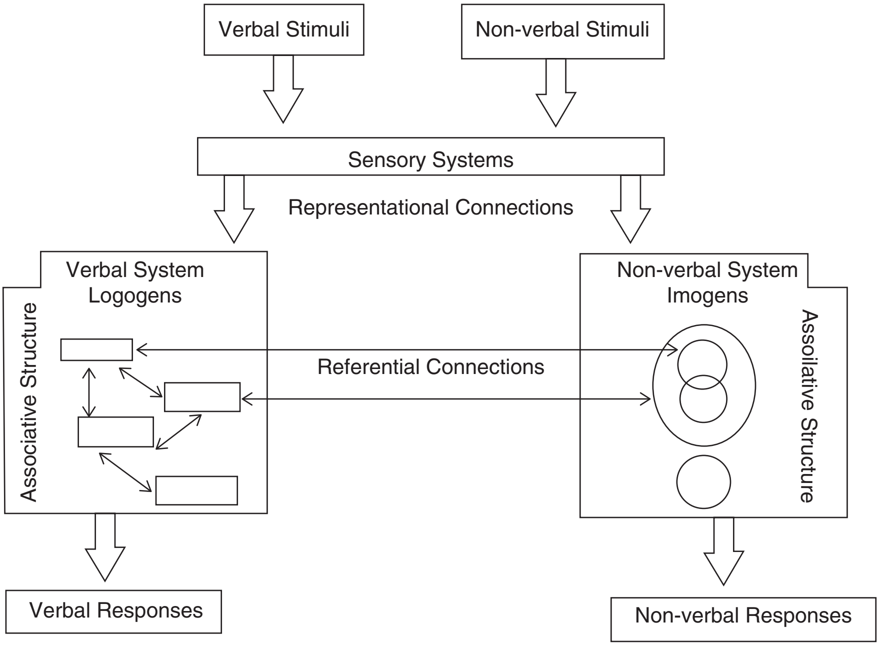


Fig. 1 Dual coding model

events and objects are controlled in a cognitive system separate from the system controlling verbal input. The sub-system handling language specific information is labelled the verbal sub-system, whilst the non-verbal system can be referred to as the symbolic or imagery sub-system, as its primary purpose is the creation of mental images and scenes (Paivio 1986). Specifically, the former sub-system contains "auditory, visual, articulatory, and other modality-specific verbal codes" (Clark & Paivio, 1991, p. 151), whereas the latter sub-system comprises "modality-specific information of shapes […], environmental sounds […], actions […], skeletal or visceral sensations related to emotions […], and other nonlinguistic objects and events" (Clark & Paivio, 1991, p. 151). These two sub-systems do not only differ with regard to the type of memorised mental codes but also with respect to the encoding method. For instance, information in the verbal sub-system can only be processed in hierarchical order and in the form of arbitrary symbols, whereas encoding in the imagery sub-system occurs simultaneously and produces denotive representations (Clark & Paivio, 1991). For example, the noun *book* is arbitrary, as different languages use distinct words to label this object, and it has to be integrated into a hierarchical, syntactic structure which is governed by rules. In contrast, the mental image of a book denotes its tactual representation and it is embedded in the mind with simultaneously constructed memories, such as the occurrence of sounds at the time of interaction with the book. However, Paivio (1986) also stresses that these two subsystems can work cooperatively. This is depicted in Fig. 1 above, which is an adaption of Clark and Paivios's (1991) illustration.

As displayed in Fig. 1, the verbal and imagery systems are connected by either associative or referential connections, meaning that activity in one system can lead to the initiation of the other (Paivio, 1986). *Logogens* refers to the units which are activated when verbal input is received, whereas *imagens* label non-verbal units (Paivio, in Heredia & Attariba, 2013, p. 43). As the term *associative connections* implies, these connections facilitate the association between either verbal or non-verbal representations (Clark & Paivio, 1991). For instance, the noun *microscope* could evoke words such as *fascinating* or *biology* in the mind. Additionally, the image of a microscope could elicit the memorisation of other visual input, such as adjusting its focus or the sound of placing a sample under its lens. With respect to referential connections it is evident in Fig. 1 that these refer to the connections between verbal and non-verbal representations. For instance, the noun *microscope* might generate the image of a microscope or the face of a laboratory partner. This cross-referencing also works vice versa, meaning that the object of a microscope evokes the noun *microscope* in the mind. Notably, DCT stresses the importance of past experiences as it proposes that mental codes are encoded with varying strengths depending on the activity level during encoding (Clark & Paivio, 1991). What is more, creating both verbal and non-verbal codes in the mind is not only crucial for individuals to comprehend their environment but also for SLA, which is why I introduced this model initially. The succeeding paragraph will therefore elucidate DCT's approach to language learning and its connection to language aptitude.

Acquiring new vocabulary is dependent on the creation of associations in LTM. DCT extends this concept and proposes that both verbal and non-verbal associations need to be created at the time of studying a foreign language in order to ensure its long-term memorisation. This is due to the fact that mental images for words can be produced as rapidly as associations between words, which indicates that imagery is highly relevant for the comprehension of both single words and whole texts (Clark & Paivio, 1991). For instance, studies show that the creation of an image in the mind when studying paired associates enhances the probability of the successful retrieval of the newly acquired word (Paivio, in Heredia & Attariba, 2013). Expressed differently, images function as additional retrieval cues to verbal information which form stronger associations between the familiar and the foreign word and hence, contribute to facilitating the retrieval of lexical input. However, DCT does not only address the beneficial effects of the creation of non-verbal associations but also emphasizes that indirect verbal associations positively contribute to long-term encoding and the retrieval process (Clark & Paivio, 1984). For instance, when studying foreign vocabulary for the words *flower*, *grass*, and *tree* the creation of an association of these expressions with *plant* can also function as a cue facilitating the retrieval process. This suggests that individuals who use dual coding techniques when studying a second language create an advantage for themselves over learners who solely focus on the verbal input. It is therefore palpable that DCT embodies an ID that is a component of language aptitude, since not all language learners use dual coding techniques similarly. For instance, the ease with which individuals create non-verbal associations varies significantly between learners (Clark & Paivio, 1991). Additionally, it was found that the ability to use images as cues is also dependent on whether the learner is a left or right hemispheric thinker. Studies reveal that the left hemisphere is specialised in verbal tasks, whereas the right hemisphere is dedicated to imaginal tasks (Mildner, 2008). This means that "[…] right hemispheric preference thinkers may be inclined to use more mental imagery" (Sadoski, 2005, p. 224), whereas individuals utilising the left hemisphere may primarily create verbal representations. These examples stress that language acquisition depends on an individual's cognitive, innate abilities, although learners can enhance their performance when directly instructed to image while studying (Clark & Paivio, 1991), meaning that environmental factors cannot be disregarded.

DCT is an older model of associative language learning but it is crucial nevertheless for language aptitude and SLA research, since the idea of separated verbal and non-verbal processing systems is also inherent in subsequent memory models. For instance, the multi-component model of WM, which will be described in the following section, also distinguishes between these two systems.

1.2 (Verbal) Working Memory and Language Aptitude

The term WM appears to have been proposed firstly in the 1960s by Miller, Galanter, and Pibram in their book *Plans and their Structure of Behaviour* and has subsequently been adopted by cognitive psychology. WM is described as a device capable of temporarily storing information, which can be manipulated in order to perform cognitive tasks, such as calculating tips or remembering phone numbers (Baddeley, 2003). Since material is lost rapidly in this memory function, WM can be characterised as a system of STM whose function is to focus attention (Baddeley, 2007) in order to "support complex cognitive activities" (Baddeley et al., 1998, p. 152). However, STM and WM are distinct as WM is capable of computing cognitive processes and of functioning independently in contrast to STM, which is merely a storage device (Wen & Skehan, 2011). What is more, WM is regarded as a central memory system when investigating language aptitude as research demonstrates (Biedron, 2012; Wen, Biedron, & Skehan, 2016). For instance, Wen and Skehan (2011, p. 21) propose that WM plays an essential role in SLA and should therefore be regarded as a component of language aptitude. Moreover, Miyake and Friedman (1998, p. 339) label WM as "one (if not the) central component of this language aptitude". These findings are unsurprising when considering the internal workings of WM, which reveal that this memory system provides all necessary components to acquire novel material such as a foreign language. Its structure can be thoroughly described with Baddeley and Hitch's multi-component model of WM from 1974 which replaced Atkinson and Shiffrin's (1968) two-component model.

The multi-component model originally consisted of three parts, namely the phonological loop, the visuospatial sketchpad, and the central executive (Baddeley, 2007). The central executive functions as an "attentional controller" (Baddeley, 2000) which is supported by both the phonological loop, which controls the processing of speech-based input, and by the visuospatial sketchpad, which is

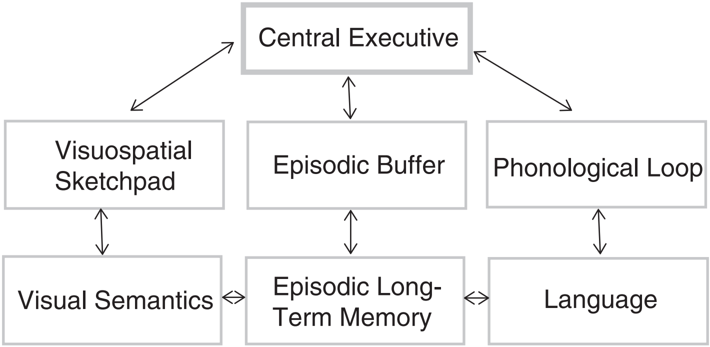


Fig. 2 Multi-component model of working memory

responsible for the transformation of visual material (Baddeley, 2000). A fourth component, the episodic buffer, was introduced in 2000 by Baddeley and Hitch in order to provide a better explanation for the creation of LTM traces, since it is crucial for "long-term episodic learning" (Baddeley, 2000, p. 421). For illustration, Fig. 2 below displays the interactions between the three slave systems (phonological loop, visuospatial sketchpad, and episodic buffer) and the central executive:

As Fig. 2 demonstrates, the central executive operates as a supervisor component, which focuses attention to control and coordinate information from the subsidiary systems (Rota & Reiterer, 2009). Although all components of WM are inter-related and contribute to the acquisition of novel information, the focus lies on the phonological loop in this paper, as it is responsible for the processing of speech sounds and regarded as the component most crucial for SLA (Baddeley, 2003). As aforementioned, previous research found significant correlations between verbal WM and language aptitude. For example, studies investigating the connections between the phonological loop and language aptitude were conducted and produced supportive findings for the proposal that these two components are highly correlated. Papagno and Vallar (1995) generated significant results demonstrating that better learners of a second language (L2) clearly outperform poorer learners in phonological memory tests. Other researchers, including Baddeley et al. (1998), endorse that verbal WM, hence the phonological loop, directly correlates with SLA and that it is therefore an integral part of language aptitude.

At this point it should be clarified that there is also a difference between verbal WM and verbal STM, albeit these two memory systems appear so similar that some researchers use the terms interchangeably (Cowan, 2008). As Verhagen and Leseman (2016) explain, verbal STM is merely capable of storing verbal information, whereas verbal WM has the capacity of manipulating verbal information while it is being stored. With this in mind, the phonological loop component needs further explanation.

The phonological loop is a device capable of temporarily storing novel phonological forms, while permanent traces are constructed in LTM (Baddeley et al., 1998). In other words, this component functions as a "language learning device"

(Baddeley et al., 1998, p. 158), since words of a foreign language embody such unfamiliar linguistic forms. Baddeley (2003) proposes that the phonological loop possibly evolved specifically for the purpose of SLA. In a series of experiments he finds that a defective phonological loop also negatively influences language learning. It is therefore palpable that verbal WM, or phonological loop capacity, is closely related to language aptitude.

The crucial role of the phonological loop results from its internal workings which operate as follows: It comprises two sub-systems, namely a rehearsal and a storage function. Information is received by the phonological store where it is forgotten unless it is repeated by the rehearsal store (Baddeley, 2003). Hence, re-articulation, either loudly or in the mind, facilitates memorisation and encoding into LTM (Rota & Reiterer, 2009). A logical consequence of the suppression of rehearsal is that novel information cannot be stored in the phonological loop. This effect is also achieved when long or similar sounding phonological material has to be remembered (Baddeley, 2000). Expressed differently, phonological loop capacity and encoding are influenced by the effects of "word-length", "phonological similarity", and "articulatory suppression" (Rota & Reiterer, 2009, p. 81). These effects need to be avoided in order to overcome decay in verbal WM. It can therefore be stated that the phonological loop stands at the beginning of a language acquisition chain, since, through the process of rehearsal, fixations of unfamiliar phonological forms in LTM are enabled.

In the previous paragraph connections to LTM are mentioned because WM and LTM are inter-reliant systems that should not be examined as isolated categories when analysing language aptitude. For instance, WM does not only generate connections to LTM but LTM can also access WM in order enhance its capacity. For example, Jones and Macken (2015, p. 2) state that "long term lexico-phonological representation" may be used to prevent the decay of representations in short-term storage. In other words, LTM can assist WM in preventing an overload and therefore enhances WM capacity (Waters, 2015). Additionally, WM and DM, which is a form LTM, are related systems supported by common brain networks (Lum, Ullman, & Conti-Ramsden, 2015). Lum et al., (2015), p. 77) describe that WM "supports the encoding of information into declarative memory by […] chunking information prior to being encoded into the hippocampus". Moreover, they also suggest that WM functions as a temporary store for controlling material retrieved from DM. Hence, due to the high inter-relatedness of WM and DM, the latter is also under consideration in this paper, and will be described in the proceeding section.

1.3 Declarative Memory and Language Aptitude

DM is an element of the declarative/procedural (DP) model, which was proposed by Michael Ullman, a professor of neurolinguistics at Georgetown University. He states that language cannot be viewed as an isolated category and hence non-language data should be considered for the investigation of language in the brain (Ullman, 2004).

Ullman therefore investigates DM and procedural memory (PM), two well-studied brain systems responsible for distinct memory functions. In DM, information regarding facts and events is memorised, whereas knowledge about cognitive skills and habits, such as driving a car, is remembered in PM (Ullman, 2001a). Considering these different types of knowledge memorised by each system, learning in DM is rapid and explicit in contrary to PM, which only allows gradual, implicit learning (Ullman, 2004). The DP model presupposes that the distinction between these two memory systems is in conjunction with the distinction between lexicon and grammar (Ullman, 2001a). Mental grammar, a system governing the rules of language, organises the mental lexicon, the words of a language, on both the syntactical and the morphological level (Ullman et al., 1997). Ullman et al. (1997) discover that temporal-parietal/medial-temporal brain systems underlie both the mental lexicon and DM, and that mental grammar and PM are mutually supported by frontal/basal-ganglia structures. This means that the brain structure controlling mental grammar and PM regulates not only cognitive skills, but also "the computation of already-learned, rule-based procedures" (Ullman, 2004, p. 245), such as forming syntax. Additionally, the brain system supporting mental lexicon and DM is not only responsible for the memorisation of facts, but also of word sounds and meanings, arbitrary relationships, and "idiosyncratic word-specific knowledge" (Ullman, 2004, p. 244), such as the association of the expression *table* with the object of a table. Even though mental grammar and PM, and mental lexicon and DM are inter-related systems, the latter stands in the centre of this paper, since this element is most crucial with regard to SLA (Ullman, 2005). Although the creation of grammatical forms depends on PM in L1, the DM system of late L2 learners is also capable of acquiring rules of a foreign language (Ullman, 2001b). This is due to the fact that the learning abilities in DM remain strong during childhood, whereas the function of PM decreases with age, and the acquisition of grammar in this system becomes difficult (Ullman, 2001b).

Although there appear to be no studies investigating this hypothesis in particular at the time of writing, Papagno and Vallar (1995) suggest that the lexical knowledge of phonological forms contributes to the acquisition of non-words. Thus, they conclude that SLA partly depends on the pre-existence of lexical knowledge, hence LTM. This ties in with LTM's ability of accessing WM in order to enhance its capacity, as described in the previous section. Similarly, Golestani and Zatorre (2008) indicate that the creation of novel, phonological categories in LTM is a crucial step towards the successful learning of speech sounds. Additionally, Ullman (2005, p. 125) states that "[m]emorizing complex forms and rules in declarative memory may be expected to lead to a fairly high degree of proficiency", which appears logical since DM in L2 learners provides the basic tools necessary for the production of a foreign language: lexicon and grammar.

1.4 Language Aptitude and Education

As aforementioned, I also take education into account with the purpose of investigating whether an environmental factor is capable of influencing both memory and language aptitude. There is already a general consensus regarding the positive effects of formal schooling on cognition, which is thought to “[…] shape life-long cognitive development” (Martensson & Lövdén, 2011, p. 1). For instance, formal schooling was found to prevent dementia (McDowell, Xi, Lindsay, & Tierney, 2007) and to potentially increase intelligence (Brinch & Galloway, 2012). Most importantly for this study, formal schooling was also found to enhance memory performance: Research demonstrates that literacy has a significant impact on WM and that formal schooling per se embodies a means of enhancing memory capacity (Kosmidis, Zafiri, & Politimou, 2011). In addition, there is evidence from previous studies that the degree of education has substantial effects on the development of mental capabilities, including memory and those factors impacting it. For instance, a study examining the effects of education on verbal span tests demonstrates a strong correlation between education and test results (Orsini et al., 1986), whereupon a higher number of schooling years resulted in higher test scores. A similar study was conducted in 2014 (Fastame, Hitchcott, & Penna), which also found that a higher degree of education positively impacts verbal memory significantly. Additionally, studies regarding the influence of education on DM were also conducted. For example, Martensson and Lövdén (2011) found that associative memory capacity, including DM capacity, can be improved by formal schooling. Although there appears to be a consensus concerning the positive influence of education on memory across a number of studies, research in this area also produces conflicting findings. For instance, Gómez-Pérez and Ostrosky-Solís (2006) assert that education does not significantly correlate with memory. Nevertheless, they do find that attention, a factor enhancing WM capacity, is influenced by education (Awh, Vogel, & Oh, 2006).

Evidently, my study is distinct to previous work in the field, since a direct correlation between education and language aptitude is assumed. Additionally, education is defined differently, as will be specified in the methodology section. Nonetheless, attention is also considered a key component which is expected to produce differing results, since a higher ability to focus attention leads to higher WM capacity (Engle, 2002), which, as described in Sect. 1.2, increases language aptitude. What is more, previous research depicts that attention-WM is the cognitive function “most sensitive to educational level” (Gómez-Pérez & Ostrosky-Solís, 2006, p. 487). As described in Sect. 1.2, WM and DM are highly inter-reliant systems and it is therefore expected that the ability to distribute attention will also influence DM. Cowen (1997), for instance, describes that the recall of explicit knowledge is only possible when attention is involved in both the encoding and the retrieval process. It is therefore expected that DM will be indirectly impacted by attention.

Based on the information provided in the previous sections, the following concrete hypotheses regarding the relationship between language aptitude and verbal WM, DM, and education can be formulated: Firstly, H1 assumes that an increase in verbal WM capacity will result in a higher language aptitude. More precisely, it is predicted that a high phonological loop capacity, hence a high verbal WM capacity, leads to rapid language learning and an increased language aptitude. Secondly, H2 infers that a higher DM capacity will increase language aptitude. Thirdly, H3 assumes that a higher level of education is an indirect predictor of a higher language aptitude, due to it increasing different memory abilities. Finally, H4 assumes that a higher amount of months spent partaking in an education is an indirect predictor of a higher language aptitude, since an increasing length of education is often represented by an increase in difficulty, which should both challenge and improve memory abilities.

2 Methodology

2.1 Participants

I present cross-sectional data from a study comprising 30 volunteering adults aged between 19 and 29 years (*M* = 22.43, *SD* = 2.445), all of whom are German native speakers born and raised in Austria, and have completed the "Matura", meaning the Austrian version of A-levels. The sample contains both male (46.7%) and female (53.3%) participants. The participants are divided into two groups differing in educational/occupational status: university students (*n* = 15, *M*(age) = 22.93, *SD*(age) = 2.939, 46.7% males, 53.3% females), and workers who have never attended a university (*n* = 15, *M*(age) = 21.93, *SD*(age) = 1.797, 46.7% males, 53.3% females). The workers' occupations lie in the areas of retail, nursing, kindergarten teaching, machine construction, gardening, waitressing, and army service.

2.2 Instruments

For the accumulation of data the digit span test from the revised version of the Wechsler Adult Intelligence Scale (Wechsler, 1939), the non-word test (Benner, 2005), the Mehrfachwahl-Wortschatz-Intelligenztest (MWT-B; Lehrl, 1989), and section V of the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 1959), were used.

The digit span test and non-word test were both used to assess verbal WM. The Wechsler Digit Span test consists of two sub-tests, namely the digit span forward test, and the digit span backward test, both of which measure the capacity of phonological input which can be remembered. For the former, the participants are required to remember and then loudly repeat strings of numbers between three and nine digits. For the latter, the subjects attempt to perform the same task in reverse, except with number strings ranging from two to eight digits. Both digit span tests assign two rows to each string of numbers. If both rows of a single digit span length are repeated incorrectly, the test is terminated.

The non-word test assesses the accuracy with which novel phonological forms are memorised. It consists of non-words which resemble German phonology, with the memorisation and recollection increasing from two to eight non-words as the test progresses. This test is also terminated under the same conditions as the Wechsler Digit Span test.

The MWT-B was used to determine DM capacity, since it is designed to measure verbal intelligence, hence the capacity of the mental lexicon. Specifically, the MWT-B is a test which assesses the vocabulary capacity participants have in their L1 (German), meaning that both implicit and explicit knowledge is taken into account. It consists of 37 rows of words, which contain 1 German word and 4 non-existent, German sounding distractors each. The participants are required to find the correct German word in each row, a task which becomes more difficult as the test progresses due to the register increasing. Participants with a higher score possess a greater mental lexicon.

The MLAT measures the ability of learning a foreign language and consists of four independent components, whereupon in this study the focus lies on the *rote learning ability of a foreign language* component only. This rote learning ability is assessed with the MLAT V, labelled *paired associates*, where participants are required to acquire pseudo-Kurdish vocabulary within 2 min, meaning that only explicit learning abilities are tested. The test originally consists of 24 Kurdish-English word pairs. However, for this test the English vocabulary was translated to German in order to avoid a distortion of results due to the native tongue of the participants. After the prescribed time frame for memorisation ends, the subjects are required to fill out a multiple choice test which provides one correct German translation and four distractors for each of the pseudo-Kurdish words. Those who score higher on the MLAT V have a greater ability of acquiring vocabulary of a foreign language.

2.3 Procedures

Participants were collected by word of mouth; 75% of the sample approached participated in the study. The purpose and nature was explained to all participants before they signed the consent form, which was translated to German in order to make the specifications clearly understandable. The data was collected in both public (university) and private (participants' homes) locations, whereupon it was ensured that each location provided a quiet space for the implementation of the study to avoid sources of distraction. The participants completed all three tests within 30–40 min approximately.

3 Results

Data was analysed through SPSS (version 21). For validation purposes, data was tested for normality across a variety of factors and groups with the Shapiro-Wilk test. The whole sample ($N = 30$) was tested for normality of distribution which demonstrated non-significance, hence a normal distribution, for both digit span ($p = 0.335$) and MLAT V ($p = 0.186$). However, normality of distribution was violated for both non-words ($p = 0.003$) and verbal IQ ($p = 0.019$) for the whole sample. The two groups, university students ($n = 15$) and workers ($n = 15$), were also tested for normality of distribution across these four categories. Results show non-significance for digit span (ps = 0.307, pw = 0.232), MLAT V (ps = 0.625, pw = 0.097), students' non-words ($p = 0.267$), and verbal IQ (ps = 0.575, pw = 0.837), meaning that a normal distribution can be assumed. However, normality of distribution was violated for workers' non-words ($p = 0.002$). Additionally, the two components education length (amount of months students spent at university) and profession length (amount of months workers spent in their profession) were tested for normality of distribution in MLAT V scores. Results show a normal distribution for both education length ($p = 0.343$) and profession length ($p = 0.075$). Education length and MLAT V scores, and profession length and MLAT V scores were also checked for linearity as this was a requirement for the multiple linear regression analysis, which will be conducted below when H4 is addressed. The scatterplots for both education length and MLAT V, and profession length and MLAT V demonstrated that linearity was not violated.

The whole sample and the two groups, students and workers, were tested with the digit span and non-word tests to assess verbal WM. DM capacity, or verbal IQ, was tested with the MWT-B. Language aptitude for rote learning was assessed with the MLAT V. An overview of the scores of the whole sample ($N = 30$), students ($n = 15$), and workers ($n = 15$) across these tests is provided in Table 1 below.

H1 assumes that a high phonological loop capacity, hence a high verbal WM capacity, leads to rapid language learning and an increased language aptitude. To test for the relationship between verbal WM and language aptitude, I tested the whole sample ($N = 30$) with a Pearson product-moment correlation between digit span and MLAT V, and non-words and MLAT V. The results are depicted in Figs. 3 and 4 below.

As Fig. 3 illustrates, there is no correlation between digit span and language aptitude as the relation is non-significant ($r = 0.077$, $p = 0.783$). This can be seen in Fig. 3 by the spread of the individual markers on the scatterplot.

As evident in Fig. 4, there is no correlation between non-words and MLAT V since the relation is insignificant ($r = 0.068$, $p = 0.720$).

As the normality of distribution was violated for non-words for the whole sample, Spearman rank-order correlation was conducted to check the results. This correlation also shows no significant relation between non-words and MLAT V scores ($r_s = 0.073$, $p = 0.702$).

Table 1 Overview scores by sample group for digit span, non-words, verbal IQ and MLAT V

Sample/test	Mean	SD	Maximum	Minimum
Digit span				
Students	15.33	3.16	21.00	11.00
Workers	13.93	3.85	23.00	9.00
Whole sample	14.36	3.53	23.00	6.00
Non-words				
Students	4.40	1.45	7.00	2.00
Workers	3.13	1.55	7.00	2.00
Whole sample	3.77	1.61	7.00	2.00
Verbal IQ				
Students	114.07	12.45	136.00	94.00
Workers	99.09	5.87	112.00	88.00
Whole sample	106.57	12.23	136.00	88.00
MLAT V				
Students	17.00	4.96	24.00	7.00
Workers	12.20	3.19	17.00	8.00
Whole sample	14.60	4.77	24.00	7.00

$N = 30$, n(students) = 15, n(workers) = 15

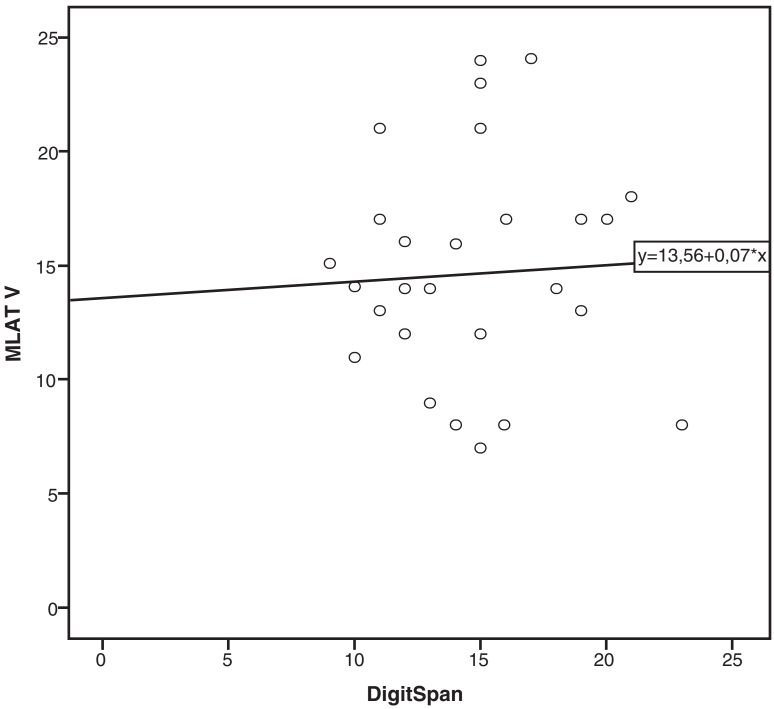


Fig. 3 Effect of digit span on MLAT V scores

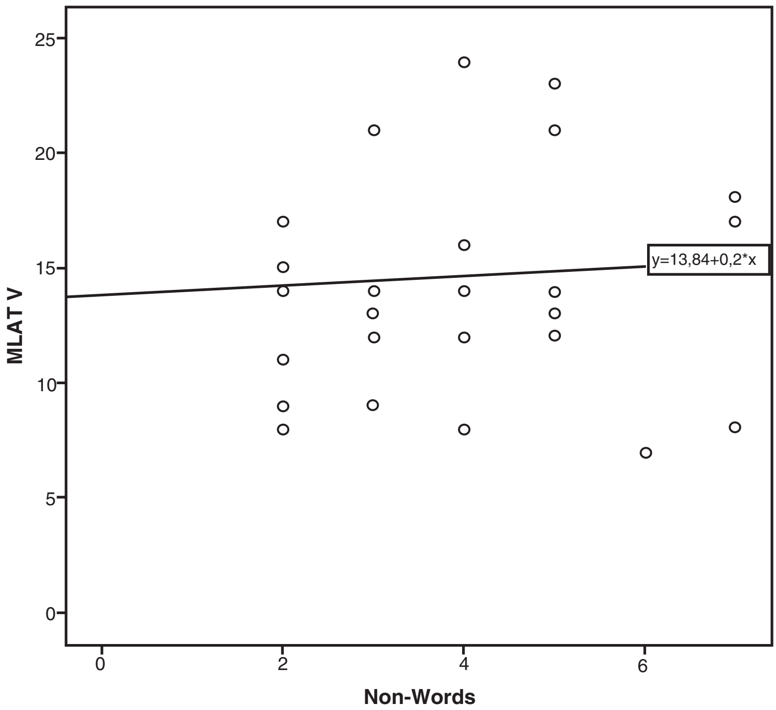


Fig. 4 Effect of non-words on MLAT V scores

A Pearson product-moment correlation was also used to test H2 which infers that a higher DM capacity will increase language aptitude. The whole sample ($N = 30$) was therefore tested for a correlation between the MWT-B and MLAT V. Figure 5 below depicts the relation between verbal IQ and language aptitude.

As depicted in Fig. 5, verbal IQ and language aptitude share a positive, moderate, linear relationship as the markers are amassed around the regression line. This finding is significant ($r = 0.43$, $p = 0.018$).

Since the normality of distribution was violated for verbal IQ for the whole sample, I checked these results with a Spearman rank-order correlation. It demonstrated a significant, positive, moderate, monotonic correlation between the two variables ($r_s = 0.417$, $p = 0.022$).

H3 assumes that a higher level of education is an indirect predictor of a higher language aptitude due to it increasing different memory abilities. To test this hypothesis, an independent samples t-test was conducted to compare the scores of non-words, digit span, verbal IQ, and MLAT V between students ($n = 15$) and workers ($n = 15$). There was a significant difference in the scores for non-words of students ($M = 4.40$, $SD = 1.454$) and workers ($M = 3.13$, $SD = 1.552$); $t(28) = -2.307$, $p = 0.029$. The difference in the scores for digit span of students ($M = 15.33$, $SD = 3.155$) and workers ($M = 13.93$, $SD = 3.845$) was not significant; $t(28) = -1.09$, $p = 0.285$. There was a significant difference in the scores for verbal IQ between

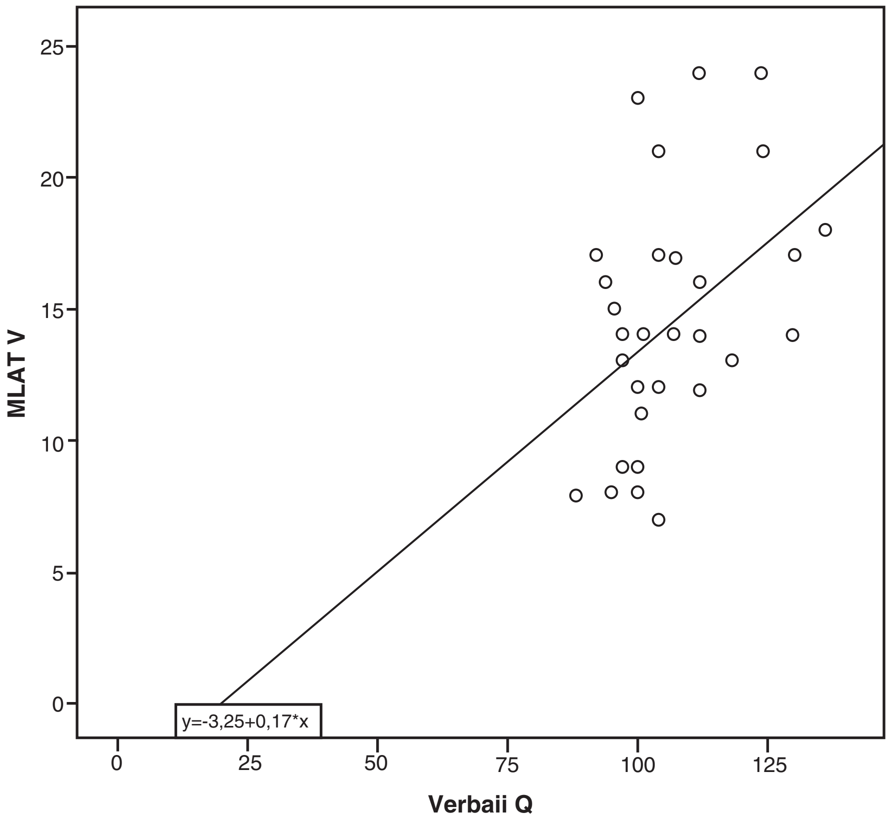


Fig. 5 Effect of verbal IQ on MLAT V scores

students (M = 114.07, SD = 12.447) and workers (M = 99.07, SD = 5.873); $t(28) = -4.221$, $p < 0.001$. Finally, there was also a significant difference in the scores of the MLAT V of students (M = 17.0, SD = 4.957) and workers (M = 12.20, SD = 3.189), $t(28) = -3.154$, $p = 0.004$.

Since the normality of distribution was violated for workers' non-words, the non-parametric Mann-Whitney U test was conducted to compare the scores of non-words between students (n = 15) and workers (n = 15) again. The results show that the null-hypothesis can be rejected, meaning that the distribution of non-words is not the same across categories of education; this is significant ($p = 0.21$).

H4 assumes that a higher amount of months spent partaking in an education is an indirect predictor of a higher language aptitude, since an increasing length of education is often represented by an increase in difficulty, which should both challenge and improve memory abilities. To test this hypothesis, a multiple linear regression was calculated to predict MLAT V scores based on education length and profession length. A significant regression equation was found ($F(2/27) = 3.430$, $p = 0.047$), with an R^2 of .203. This means that 20.3% of the variance in the MLAT V scores is explained by the impact of education and profession length. Participants predicted that scores of the MLAT V are equal to 13.103 + 0.08 (education length) – 0.015 (profession length), where education length is coded as the amount of months students spent at university, and profession length as the amount of months workers spent in their profession. Participants' MLAT V scores increased by 0.08 points for each month spent at university, and decreased by 0.015 points for each month spent in a profession. Only education length was a significant predictor for participants' MLAT V scores ($p = 0.048$); profession length was an insignificant predictor ($p = 0.818$).

4 Discussion

The purpose of this study was to investigate the influence of WM, DM, and education on language aptitude, and whether there is a relationship between the cognitive factor and the environmental component. H1, which assumed that an increase in verbal WM capacity will result in a higher language aptitude, failed to be supported. The correlation analyses revealed no significant links between neither digit span and language aptitude, nor non-word span and language aptitude. This finding contradicts expectations, since the MLAT V tests rote learning ability, meaning that vocabulary has to be acquired explicitly, which should heavily involve the phonological loop since it is responsible for processing speech sounds (Baddeley, 2003). An explanation for such contradictory results might lie in the research design, since the MLAT V does not consider the production of mental images as retrieval cues in addition to mental verbal codes. It is not only crucial to divide the sub-components of language aptitude regarding phonological or grammatical abilities but to also consider the distinct memory systems involved in language acquisition. In this paper it was expected that verbal WM is the central memory system for both non-word and digit span tests and the MLAT V. This means that the encoding of auditory information was in the focus of this study and that the processing of information was therefore expected to solely depend on the phonological loop. These assumptions embody a flaw in the study, as the MLAT V tests for paired associates with one word being a concrete expression in the participants' mother tongue. As described in DCT, when learning novel vocabulary referential connections are formed in associative memory, meaning that the verbal information is linked to a mental image in the mind in order to facilitate encoding and retrieval (Clark & Paivio, 1991). This means that participants who study the German translation of the pseudo-Kurdish word *kete*, which is *Kamel* in German (or *camel* in English), may not only use the rehearsal technique to encode the verbal input but also form the image of a camel in their mind to aid them later in the retrieval process. Thus, the non-significant results of the correlation analysis may have occurred due to the fact that the digit span and non-word tests heavily rely on phonological loop performance, whereas the MLAT V is co-dependent on both auditory and visual information. This explanation appears feasible, as other research in the field stresses that it is essential to differentiate between the acquisition of word associations and phonological word representations. For instance, Duyck, Szmalec, Kemps, and Vandierendonck (2003) found that verbal WM is only crucial for the acquisition of phonological information and its LTM encoding, however associations are learned through other systems which are capable of generating mental images, such as the visuospatial sketchpad.

An additional reason for insignificant findings due to research design may be that the MLAT was not used in its complete form. The full test battery consists of five parts, namely *number learning*, *phonetic script*, *spelling cues*, *words in sentences*, and *paired associates*. However, this study only tested for *paired associates* which might have caused discrepancies leading to a non-significant result, as this section alone does not effectively measure all the components that language aptitude comprises. For instance, as the parts *number learning* and *phonetic script* both depend on auditory processing rather than associative processing, the probability of finding correlations between these parts and verbal WM may have been higher.

H2, which assumed that a higher DM capacity will increase language aptitude, was significantly supported as the bivariate correlation analyses demonstrated. Although this hypothesis appears to not have been investigated, it was not unexpected to find supportive results. This is due to the fact that research in this field assumes a strong correlation between DM and language proficiency (Ullman, 2005). Furthermore, studies also demonstrate that the pre-existence of lexical knowledge is crucial for SLA (Papagno & Vallar, 1995) and that the successful acquisition of speech sounds is dependent on the creation of phonological representations in LTM (Golestani & Zatorre, 2008). On account of these supportive findings of the strong influence of the mental lexicon on L2, it is not surprising that a significant relation between DM and language aptitude was found. In addition, as aforementioned, the MLAT V tests for *paired associates*, meaning that not only auditory but also visual information is encoded as retrieval cues in order to form associations in memory. Moreover, as Hof and Mobbs (2010) elucidate, DM facilitates vocabulary acquisition and contains the mental lexicon, which can be described as the memory system consisting of a plethora of associative, arbitrary sound-meaning relationships (Breitenstein et al., 2005). Hence, both the MLAT V and the MWT-B rely not only on phonological representations but also on mental associations, which also explains why a correlation was found.

H3, which expected that a higher level of education is an indirect predictor of a higher language aptitude due to it increasing different memory abilities, was partly supported, as evident in the parametric and non-parametric independent samples t-tests. These demonstrate that students significantly outperformed workers in the non-words, verbal IQ, and MLAT V tests, however not in the digit span test.

Even though no studies investigating this issue in particular appear to currently exist, there are similar studies investigating the relationship between memory and education. For instance, studies found significant results that a higher education increases verbal memory (Fastame et al., 2014; Orsini et al., 1986). This is also partly supported by the current study, as students outperformed workers significantly in a part of the verbal WM test, namely in the non-word test. What is more, students outperformed workers in verbal IQ, meaning that they possess a richer mental lexicon and hence, a higher DM capacity. Thus, this study supports the view that education impacts memory and stands therefore in contrast to other studies which found no correlation between these two variables (Gómez-Pérez & Ostrosky-Solís, 2006). In this study it was found that memory can be viewed as the mediating variable linking education and language aptitude. Additionally, the students outperformed the workers in the MLAT V, which can also be explained with students' educational background. As mentioned earlier, the MLAT V tests not only for verbal WM but also involves associative memory capacity and as Martensson and Lövdén (2011) found, associative memory can be improved with formal schooling. Hence, students are thought to have an advantage over workers due to their constant exposure to formal education.

Further explanations for such significant differences between university students and workers can be found when investigating another mental capability, namely attention distribution. As students are frequently exposed to a plethora of complex material, which they are required to comprehend and remember, they are thought to be capable of distributing their attention better than workers, whose daily tasks are rather repetitive. As students are able to focus their attention better, they have an advantage when doing both MLAT V and DM tasks, as the retrieval of words is only possible when enough attention was distributed at the moment of encoding (Cowen, 1997). However, the attention variable does not offer an equally satisfying explanation when turning to verbal WM. Although WM was found to heavily depend on the distribution of attention (Engle, 2002; Gómez-Pérez & Ostrosky-Solís, 2006), university students did not outperform workers in all verbal WM tasks. While students did exceed in the non-word test, which can be explained with the attention variable, this was not the case for the digit span test, as there were no significant differences between the two groups. This might be due to the different types of verbal material that are required to be remembered in each test. While digits are frequently encountered, as opposed to non-words, they are remembered with less difficulty, since LTM can interfere with WM in order to support its capacity (Jones & Macken, 2015). However, when phonologically novel material has to be remembered the listener is required to solely rely on the phonological loop (Baddeley et al., 1998), which requires a higher focus of attention. Therefore, students had an advantage for the non-word test, whereas the conditions were the same for both students and workers for the digit span test.

H4 was also supported. It assumed that a higher amount of months spent partaking in an education is an indirect predictor of a higher language aptitude, since an increasing length of education often is represented by an increase in difficulty, which should both challenge and improve memory abilities. As the regression analysis demonstrates, students score better on the MLAT V the longer they are at university, which is a significant finding. It also shows that 20.3% of the variance in the MLAT V can be explained with education length. As there appear to be no studies investigating the effect of the length of university education on memory or language aptitude, it proves difficult to place these findings within context. However, a longer time at university might enhance memory, and thus language aptitude, because cognitive functions can improve when enough training is provided. For instance, research has found significant findings that adults can improve their fluid intelligence (Gf), that is their ability to solve novel problems, by training their WM (Jaeggi, Buschkuhel, Jonides, & Perrig, 2008). Although the WM and Gf tests are distinct, enough training in the former can improve the results in the latter. Therefore, it can be suggested that sufficient training of both WM and DM, by being exposed to complex material in lectures or seminars, can lead to an extension of memory abilities and hence, an increase in language aptitude. As students outperformed workers in the MWT-B and the non-word test, this explanation appears tangible. Although workers scored significantly lower than students on memory tests, the findings demonstrate that a longer period of time spent in their profession does not significantly affect their MLAT V scores. It appears that the length a worker spends in their occupation does not impact their ability to successfully acquire a new language. As there are no studies investigating this issue in particular, it is challenging to draw a scientifically supported conclusion. It can be speculated that university becomes more demanding the longer a course lasts, whereas the type of work the participants of this study perform essentially comprises similar tasks. Therefore, the number of months spent at work does not significantly impact language aptitude, as the tasks become neither considerably more challenging, nor less strenuous. On the other hand, the number of months spent at university does significantly improve language aptitude, as course content usually becomes more complex over time.

5 Conclusion

Although this study produced significant findings and was able to shed light on the relationships language aptitude, memory, and education share, the current research has its limitations. Firstly, it is questionable whether the results of the independent samples t-tests and the regression analysis can be held accountable, since each group of participants consisted of merely 15 members. In order to create substantial results, each group should have consisted of at least 30 members so that assumptions regarding the general population can be drawn. It is possible that the nonsignificant results regarding group differences in the digit span test and the regression analysis occurred due to the limited sample size. However, even if the sample size was sufficient, other limitations would also be encountered.

For instance, not the entire test battery of the MLAT was used. In order to make substantial propositions about language aptitude in general, the MLAT parts I–V should have been taken into account. Using only part V, *paired associates*, is a potential flaw in the study as it does not aptly measure all the components of language aptitude, rather just a portion. As aforementioned, the MLAT V does not only test for auditory but also for visual processing in contrary to the non-word and digit span tests, which primarily rely on the phonological loop. This discrepancy between these types of vocabulary learning may therefore have contributed to the insignificant results concerning H1.

In addition, due to the scope of this paper, it was not possible to include additional tests or questionnaires providing information about whether participants learn vocabulary through the phonological loop, hence on a verbal basis, or with associations, meaning with the additional support of non-verbal cues. These additional measurements could have provided substantial evidence for the explanation as to why H1 could not be supported. For instance, if the majority of participants could be revealed as right hemispheric preference thinkers, who acquire novel vocabulary with imagery associations, the answer explaining the outcome of H1 could be scientifically supported.

Finally, when measuring language aptitude and memory, group differences stood in the focus, whereupon the full extent of IDs was not considered. This is a limitation as the students may have performed better not because of their educational status but because of their IDs. A control measurement should have been implemented to account for this, such as using a test-retest design, through which the workers and students would have had their language aptitude and memory ability tested again after being exposed to complex materials over a period of time. If learning intricate contents did influence memory, and language aptitude, significantly, the students should have still improved more than workers in their second attempt. Such an additional test would have facilitated result interpretation and would have been beneficial for a further clarification of the findings.

In conclusion, this study produced unexpected findings contradicting current research in the field of language aptitude and verbal WM, since the bivariate correlation analysis revealed no significant links between these two variables. However, the research in this paper shed light on the influences of DM on language aptitude and revealed a significant relationship between these two components. What is more, both of these results emphasize the importance of associative memory for language aptitude research for two reasons: Firstly, because DM is a substantial part of the associative memory system, as the mental lexicon facilitates the processing of verbal and non-verbal information in order to create LTM traces of foreign vocabulary (Breitenstein et al., 2005). The research design of this study may have demonstrated that a higher DM capacity enhances language aptitude due to individuals' mental capabilities of creating associative memories between verbal and non-verbal information when studying new words. Secondly, because the insignificant results regarding verbal WM and language aptitude may have occurred due to the flaw that associative memory was not considered, the phonological loop alone cannot be viewed as the "language learning device" (Baddeley et al., 1998). Additionally, these findings assist in the clarification of whether there are influences of education on memory and language aptitude, since significant relations between these variables were obtained. As it was revealed in this paper, both verbal WM and DM function as mediator variables between education and language aptitude, which lead to better language aptitude test scores in participants with a higher level of education. What is more, this study investigated a novel component influencing language aptitude, namely the length of exposure to education which increases in complexity. This component is novel in the sense that previous research appears to only consider formal schooling or literacy in general with a focus on differences between individuals in the same educational system. For instance, Fastame et al. (2014) compare students with 1–9 schooling years to students with more than 9 years of formal schooling. In addition, older studies such as the one conducted by Orsini et al. (1986) also stay within the same system of education, since they divide their subjects into individuals with formal schooling experiences of 5, 6–12, and more than 12 years. However, this study extends this concept and compares two groups from distinct educational systems, namely workers without an academic background and university students. The multiple linear regression analysis demonstrated that students' MLAT V scores increased with each month spent at university, whereas the amount of months workers spent in their profession did not have an effect on their MLAT V scores. Based on the assumption that course content at university become more complex the longer a course lasts, it can be assumed that students' memory is trained more frequently than workers' whose tasks are rather repetitive.

However, considering the limitations of this study, future research is required in order to further test the findings in this paper. For instance, as H1's conflicting findings may have been produced by the research design, future research should use the entire MLAT test battery to investigate this hypothesis and also additional associative memory tests. For instance, in order to determine whether mental images were produced as retrieval cues participants should also attempt an associative memory test or fill out a questionnaire determining whether they are left or right hemispheric thinkers.

What is more, as H2, H3, and H4 have not been scrutinised in their particular form yet, more research in these areas is required in order to further clarify the relationships between language aptitude and DM, the level of education, and the length of education. I suggest that future research should attempt a study similar to this one, except that a greater sample size should be collected. Moreover, as aforementioned, additional tests examining the effect of the length of exposure to complex materials, such as lecture content, on language aptitude are required in order to facilitate a more reliable interpretation of results. This is due to the fact that this study could only draw assumptions due to a lack of research in the fields of language aptitude and education. For instance, a group of workers could be given lecture material which they are required to comprehend and remember over a specific period of time. The MLAT should be attempted by this group before and after the exposure to the intricate readings in order to further test whether education (or university in particular) has a significant impact on language aptitude. This is, of course, only one of a plethora of options to analyse the impact of education on memory and language aptitude. Other research could also contain the testing and comparing of first semester and PhD students, or of university and high school students. Moreover, a longitudinal study in which individuals are tested over their lifetime after new educational milestones were reached would also be highly beneficial for the examination of the impact of education as an environmental factor on cognitive processes.

Although this study has its flaws, I do believe that its findings shed light on the influences of the ID *memory* and the environmental factor *education* on language aptitude. With regard to my research question I could successfully demonstrate that both DM capacity and the degree of education have significant effects on language aptitude, and that the cognitive factors WM and DM are sensitive to education. Even though no significant correlations between verbal WM and language aptitude was found, this study found a palpable explanation for this outcome, which can aid future researchers in designing a research model that is better suitable for the investigation of H1 in order to produce potentially significant findings. On a final note I want to mention that this study demonstrated that the environmental factor *education* should be taken into consideration more frequently in language aptitude research, as this small scaled study was capable of producing novel, significant findings in this field. It would therefore be interesting to see results from large scaled studies. Furthermore, it might be worthwhile to investigate whether there are differences between students and workers with regard to knowledge that is solely acquired implicitly. Explicit learning, which was the central learning strategy in this paper, is influenced more by attention and exposure time than implicit learning (Malhotra, 2004), meaning that implicit cognitive abilities of students and workers might not differ to such a great extent.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation, 2*, 89–195.

Awh, E., Vogel, E. K., & Oh, S. H. (2006). Interactions between working memory and attention. *Neuroscience, 139*(1), 201–208. https://doi.org/10.1016/j.neuroscience.2005.08.023

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*, 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2007). *Working memory, thought and action*. Oxford, UK: Oxford University Press.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158–173.

Benner, U. (2005). *Syllables in speech production – A study of the mental syllabary*. Linguistic Master thesis at the University of Stuttgart, Stuttgart.

Biedron, A. (2012). Working memory and short-term memory abilities in accomplished multilinguals. *The Modern Language Journal, 96*, 290–306. https://doi.org/10.2307/41684074

Breitenstein, C., Jansen, A., Deppe, M., Foerster, A., Sommer, J., Wolbers, T., & Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *NeuroImage, 25*(3), 958–968. https://doi.org/10.1016/j.neuroimage.2004.12.019

Brinch, C., & Galloway, T. (2012). Schooling in adolescence raises IQ scores. *PNAS, 109*(2), 425–430. https://doi.org/10.1073/pnas.1106077109

Carroll, J. B. (1971). *Implications of aptitude test research and psycholinguistic theory for foreign language teaching*. Paper presented at 17th International Congress, International Association of Applied Psychology, Liege, Belgium, 1–15.

Carroll, J. B., & Sapon, S. (1959). The modern languages aptitude test. In Z. Dörnyei (Ed.). (2006). Individual differences in second language acquisition. *AILA Review, 19*, 42–68.

Clark, J., & Paivio, A. (1984). *Associative mechanisms in cognition*. Paper presented at the annual meeting of the Canadian Psychology Association, Ottawa, Canada.

Clark, J., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149–210.

Cowan, N. (2008). What are the differences between long-term, short-term, and working meory? *Progress in Brain Research, 169*(7), 323–388. https://doi.org/10.1016/S0079-6123(07)00020-9

Cowen, N. (1997). *Attention and memory*. New York/Oxford, NY: Oxford University Press.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*(1), 42–68. https://doi.org/10.1075/aila.19.05dor

Duyck, W., Szmalec, A., Kemps, E., & Vandierendonck, A. (2003). Verbal working memory is involved in associative word learning unless visual codes are available. *Journal of Memory and Language, 48*(3), 527–541. https://doi.org/10.1016/S0749-596X(02)00533-8

Engle, W. R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11*(1), 19–23. https://doi.org/10.1111/1467-8721.00160

Fastame, M. C., Hitchcott, P. K., & Penna, M. P. (2014). Does education influence visuo-spatial and verbal immediate serial recall in healthy older adults? *Quality and Quantity, 49*(5), 2155–2167. https://doi.org/10.1007/s11135-014-0099-3

Golestani, N., & Zatorre, J. R. (2008). Individual differences in the acquisition of second language phonology. *Brain and Language, 109*(2–3), 55–67. https://doi.org/10.1016/j.bandl.2008.01.005

Gómez-Pérez, E., & Ostrosky-Solís, F. (2006). Attention and memory evaluation across the life span: Heterogeneous effects of age and education. *Journal of Clinical and Experimental Neuropsychology, 28*(4), 477–494. https://doi.org/10.1080/13803390590949296

Granena, G., & Long, M. (Eds.). (2013). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam/Philadelphia: John Benjamins Publishing.

Heredia, R., & Attariba, J. (Eds.). (2013). *Foundations of bilingual memory*. New York: Springer.

Hof, P., & Mobbs, C. (Eds.). (2010). *Handbook of the neuroscience of aging*. London: Academic.

Jaeggi, S., Buschkuhel, M., Jonides, J., & Perrig, W. (2008). Improving fluid intelligence with training on working memory. *PNAS, 105*(19), 6829–6833. https://doi.org/10.1073/pnas.0801268105

Jilka, M. (2009). Assessment of phonetic ability. In G. Dogil & M. S. Reiterer (Eds.), *Language talent and the brain*. Stuttgart/Tübingen, Germany: Mouton de Gruyter.

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition, 144*, 1–13. https://doi.org/10.1016/j.cognition.2015.07.009

Kosmidis, H. M., Zafiri, M., & Politimou, N. (2011). Literacy versus formal schooling: Influence on working memory. *Archives of Clinical Neuropsychology, 26*(7), 575–582. https://doi.org/10.1093/arclin/acr063

Lehrl, S. (1989). *Mehrfachwahl-Wortschatz-Intelligenztest: MWT-B*. Perimed-Fachbuch-Verlag-Ges.

Lum, J., Ullman, M., & Conti-Ramsden, G. (2015). Verbal declarative memory impairments in specific language impairment are related to working memory deficits. *Brain and Language, 142*, 76–85. https://doi.org/10.1016/j.bandl.2015.01.008

Malhotra, K. N. (Ed.). (2004). *Review of marketing research*. New York: M. E. Sharpe.

Martensson, J., & Lövdén, M. (2011). Do intensive studies of a foreign language improve associative memory performance? *Frontiers in Psychology, 2*(12), 1–7. https://doi.org/10.3389/fpsyg.2011.00012

McDowell, I., Xi, G., Lindsay, J., & Tierney, M. (2007). Mapping the connections between education and dementia. *Journal of Clinical and Experimental Neuropsychology, 29*(2), 127–141.

Mildner, V. (2008). *The cognitive neuroscience of human cognition*. New York/London: Taylor & Francis Group.

Miyake, A., & Friedman, D. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Lawrence Erlbaum Associates.

Orsini, A., Chiacchio, L., Cinque, M., Cocchiario, C., Schiappa, O., & Grossi, D. (1986). Effects of age, education and sex on two tests of immediate memory: A study of normal subjects from 20 to 99 years of age. *Perceptual and Motor Skills, 63*, 727–737.

Paivio, A. (1971). *Imagery and verbal processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology, 38A*, 98–107.

Reed, S. (2010). *Cognition: Theories and application*. Belmont, CA: Wadsworth Cengage Learning.

Rota, G., & Reiterer, M. S. (2009). Cognitive aspects of langauge talent. In G. Dogil & M. S. Reiterer (Eds.), *Language talent and the brain*. Stuttgart, Germany: Mouton de Gruyter.

Sadoski, M. (2005). A dual view of vocabulary learning. *Reading & Writing Quarterly, 21*(3), 221–238.

Ullman, M. (2001a). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*(1), 37–69. https://doi.org/10.1023/A:1005204207369

Ullman, M. (2001b). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*(1), 105–122.

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*(1–2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 141–178). Washington, DC: Georgetown University Press.

Ullman, M., Corking, S., Coppola, M., Hickok, G., Growdon, J., Koroshetz, W., & Pinker, S. (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9*(2), 266–276.

Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology, 141*, 65–82. https://doi.org/10.1016/j.jecp.2015.06.015

Waters, A. (2015). Cognitive architecture and the learning of language knowledge. *System, 53*, 141–147. https://doi.org/10.1016/j.system.2015.07.004

Wechsler, D. (1939). *The measurement of adult intelligence*. Baltimore: Williams & Wilkins.

Wen, Z., Biedron, A., & Skehan, P. (2016). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching, 50*(1), 1–31. https://doi.org/10.1017/S0261444816000276

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting a case for "working memory as language aptitude". *Revista Ilha do Desterro, 60*, 15–43. https://doi.org/10.5007/2175-8026.2011n60p015

Aptitude for Vocabulary Acquisition

Hannah Hackl

Abstract Gathercole and Baddeley reviewed manifold research regarding vocabulary acquisition, concluding that new word learning is strongly related to phonological memory skills. According to the working memory (WM) model by Baddeley these skills are represented in the phonological loop, which is a part of WM. Consequently, it is not surprising that WM is considered to be highly significant in the process of vocabulary acquisition. To contribute to this field, the purpose of the current study is to provide further evidence underlining the importance of WM in vocabulary learning. It is hypothesized that individuals with a high working memory capacity have a higher aptitude for vocabulary acquisition. Furthermore, this study seeks to ascertain whether there are any differences between language and non-language students. Ten language students and ten non-language students participated in the study. Participants were required to complete a short questionnaire asking for personal data and were tested on their working memory span and language aptitude for vocabulary learning, measured by the Modern Language Aptitude Test (MLAT). The results indicate that language students score slightly higher in both tests. Due to these results it can be suggested that language students might be better at rehearsing, as they are constantly engaged with foreign word learning, which results in a higher WM capacity and better skills in vocabulary learning.

1 Introduction

Vocabulary acquisition is essential for acquiring a native language and learning a foreign language (Morra & Camba, 2009) and has been subject of extensive research in various ways. Researchers studied individual differences regarding young children's acquisition of words in their native language or the acquisition of words in a

H. Hackl (✉)
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria
e-mail: a1208109@unet.univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_3

foreign language of older children and adults (Morra & Camba, 2009), to name only a few examples. In 1998 Baddeley, Gathercole and Papagno proposed the most influential theory of vocabulary learning (Baddeley, Gathercole, & Papagno, 1998). They claimed that phonological short-term memory plays an important role in constructing representations of the phonological form of new words. Hence, it assists vocabulary acquisition. By converging evidence from the review of manifold research, Gathercole and Baddeley (1993) generalised that new word learning is related to "phonological memory skills". Baddeley et al. (1998) often use this term in an interchangeable way with terms such as "phonological loop capacity", "phonological loop abilities" and "phonological storage". Some of them are closely linked to the working memory model proposed by Baddeley (as cited in Morra & Camba, 2009). Hence, it is not surprising that working memory (WM) is claimed to play a crucial role for language comprehension, particularly for vocabulary acquisition. Many previous studies have already supported this assumption (Morra & Camba, 2009; Verhagen & Leseman, 2016).

Both experience and studies show that individuals learn a second or foreign language at a different pace and with a different effort, which leads to distinct outcomes. One perspective on this phenomenon is the concept of Foreign Language Aptitude (FLA) (Wen, 2012). Language aptitude is defined as a "specific talent for learning […] languages which exhibits considerable variation between learners" (Dörnyei & Skehan, 2003, p. 590). However, in contrast to other talents such as musical talent, the talent for foreign language learning consists of various independent linguistic skills (Dewaele, 2013). Carroll (as cited in Dörnyei & Skehan, 2003), for example, identifies four different components: Phonemic coding ability, grammatical sensitivity, inductive language learning ability and associative memory. Furthermore, language aptitude results from various seemingly unrelated cognitive factors that interact and determine a learner's capacity to master a second language (Dörnyei, 2006). Dörnyei (2006) adds that language aptitude "does not predict whether or not a person is able to learn a second language", but that it rather predicts "the rate of progress the individual is likely to make in learning" (p. 43) under optimal conditions. Robinson (2002) suggests that second language learning aptitude is a "complex and dynamic construct where clusters of learner variables interact with a range of L2 learning tasks and teaching techniques" (p. 113). One of these learner variables is WM, which is, in relationship with learning, regarded as "one of the most promising current directions in language aptitude studies" (Dörnyei, 2005, p. 55).

The term "working memory" can be defined as the temporary storage and manipulation of newly acquired information (Rota & Reiterer, 2009). It is assumed to be necessary for various complex cognitive activities (Baddeley, 2003), such as language comprehension, i.e. to retain earlier parts of a spoken message until they can be put into context with later parts of the message and arithmetic, i.e. to be able to maintain partial results until the rest of the answer can be calculated (Cowan, 2005). It is also necessary for drawing a conclusion. In this context WM is used to retain the thoughts while working with them. Moreover, WM is not only needed to maintain new information, but also to put it into context with pre-existing information (Cowan, 2005).

Currently, there are two main competing conceptions regarding WM (Towell, 2013). Miyake and Shah (1999) define WM as "those mechanisms or processes that are involved in the control, regulation, and active maintenance of task-relevant information in the service of complex cognition" (p. 450). This definition seems to encompass these two competing conceptions of WM's internal structure. The Multi Component Model, which was proposed by Baddeley in 1974 has more influence than the other conception and has been continuously adapted over the years. It describes WM as a separated component from the memory system in the mind (Towell, 2013). The other alternative conception of WM "relies on spreading activation" (Towell, 2013, p. 129). In contrast to the Multi Component Model it does not require WM to be a specific, separable unit (Towell, 2013). Anderson (1993) explains WM as an "expository convenience" (p. 20), which describes the currently activated part of long-term memory. As the current study is based on the Multi Component Model by Baddeley and Hitch, this model will be further described in the following paragraphs.

Three components of WM have been identified. The central executive regulates the information flow within WM, retrieves information from other memory systems such as long-term memory, and processes and stores information (Gathercole & Baddeley, 1993). Gathercole and Baddeley (1993) argue that processing resources, which are used by the central executive to fulfil these different functions, have a finite capacity. In other words, the efficiency with which the central executive performs a certain function depends on whether other demands are simultaneously placed on it or not.

The central executive is supplemented by two components, the so-called "slave systems" (Gathercole & Baddeley, 1993, p. 4). Both components are responsible for processing and temporarily maintaining material within a certain domain. The phonological loop maintains verbally coded information, whereas the visuo-spatial sketchpad is responsible for short-term processing and maintenance of visual or spatial material (Gathercole & Baddeley, 1993).

As the phonological loop is specialised in the storage of verbal material, and consequently plays an important role for vocabulary acquisition, the following subsection provides further details on this component (Fig. 1).

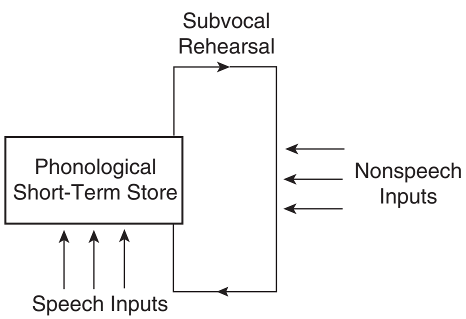


Fig. 1 Representation of the Phonological Loop, with the Phonological Store and the Articulatory Loop (Baddeley, 1986)

The phonological loop is the first and most studied component of the working memory model (Repovs & Baddeley, 2006). It consists of two subcomponents, the phonological store and the subvocal rehearsal system (Gathercole & Baddeley, 1993). The phonological store keeps memory traces for approximately 2 s in a phonological code before they fade, unless they are refreshed by the second component, the subvocal rehearsal system (Baddeley, 2003). This rehearsal system comprises a process in which memory traces are rehearsed, either verbally or sub-vocally. In this way it helps to refresh the traces and prevents them from decaying (Baddeley, 2010). However, the subvocal rehearsal system serves not only to maintain information within the phonological store, but also to register visual information in the store (Baddeley, 2003). This means that if a person is shown a sequence of letters for immediate recall, then, despite their visual presentation, the subject will subvocalise them. Consequently, the retention of the letters will depend on their acoustic or phonological characteristics (Baddeley, 2003). This theoretical concept of the phonological loop demonstrates the importance of rehearsal in vocabulary learning. Research further supports the effectiveness of rehearsing (e.g. Dahlen & Caldwell-Harris, 2013).

The concept of WM described above has proved to be durable since its publication in 1974. However, there have always been various phenomena that could not be explained by the model (Baddeley, 2000). These phenomena can be summarised in two deficits within the model (Baddeley, 2003). The first one was "a need for a system that would allow visual and verbal codes to be combined and linked to multidimensional representations in long-term memory" (Baddeley, 2003, p. 202). The second deficit can be described as "the need for the temporary storage of material in quantities that seemed clearly to exceed the capacity of either the verbal or visuospatial peripheral subsystems" (Baddeley, 2003, p. 202). Due to the deficits within the original tripartite working memory model, a new one has been created (Fig. 2). This reformulation included the proposal of a new component of WM, namely the 'episodic buffer' (Baddeley, 2003). The episodic buffer is a limited-capacity temporary storage system that is capable of retrieving information from a variety of sources (Baddeley, 2000).

The episodic buffer is assumed to be controlled by the central executive. Furthermore, the schematic representation exemplifies that the episodic buffer serves as a temporary interface between the two slave systems, the visuospatial sketchpad and the phonological loop. Due to the fact that it serves as an interface between the systems, which involve a different set of codes, the episodic buffer is assumed to use a common multi-dimensional code. Although the buffer is separated from long-term memory, it forms an important stage in long-term episodic learning (Baddeley, 2000). Juffs and Harrington (2011) describe working memory capacity as a "bottle-neck through which information has to pass in order to be permanently stored in long-term memory" (p. 139).

Many previous studies support the importance of working memory capacity in regard to vocabulary acquisition. Verhagen and Leseman (2016), for example, addressed the question of how verbal short-term memory (VSTM) and verbal working memory (VWM) relate to the acquisition of vocabulary and grammar in 5-year

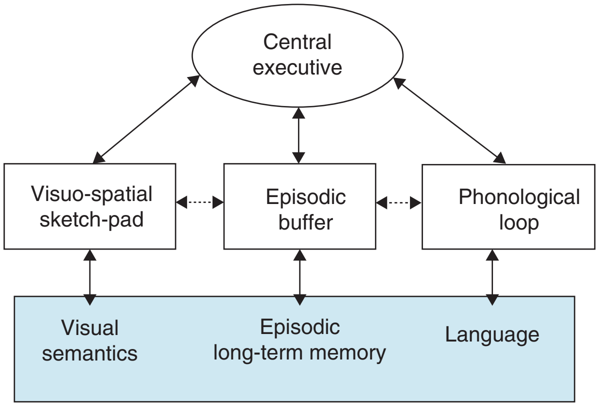


Fig. 2 Later Development of the Multicomponent Working Memory Model (Baddeley, 2010)

old L1 and L2 children in a naturalistic setting. Therefore, they tested the two components of verbal memory separately. Their subjects' VSTM was investigated, which is defined as the capacity to store verbal information, and their VWM was also tested, which is described as the ability to process verbal information while it is being stored (Verhagen & Leseman, 2016). VSTM is seen as crucial for the development of stable phonological representations in long-term memory that is needed for vocabulary and grammar learning in L1 and L2 (Baddeley et al., 1998; Speidel, 1989; Verhagen & Leseman, 2016). In contrast, VWM is argued to be important only for grammar learning (Verhagen & Leseman, 2016). The study of Verhagen and Leseman (2016) supports these assumptions, as they found out that VSTM predicted both vocabulary and grammar learning in L2, whereas VWM only predicted grammar learning. In addition, the results showed that VSTM and VWM affect L1 and L2 similarly (Verhagen & Leseman, 2016). A study by Chrysochoou, Bablekou, Masoura, and Tsigilis (2013) examined the relations between Greek vocabulary knowledge and VWM and VSTM measures in children aged between 5.5 and 9.5 years. The preschool and elementary school years are periods of remarkable changes regarding vocabulary and WM (Chrysochoou et al. 2013; Cowan, 1997; Pickering & Gathercole, 2001). The results showed that VSTM and VWM affect vocabulary development in the early years. However, the relations between VSTM, VWM, and vocabulary acquisition declined with age (Chrysochoou et al., 2013).

WM does not only influence vocabulary development. As already mentioned, it also plays a key role in the language aptitude construct. In a study carried out by Yoshimura (2001) WM scores correlated significantly with overall language aptitude scores. By splitting the overall scores into its components, it was shown that WM correlated only with the word associates and language analysis components; however, the sound-symbol components did not show a significant correlation with WM. Yalçın, Çeçen, and Erçetin (2016) conducted a similar study: They examined whether "WM capacity [is] significantly related to both overall language aptitude scores and subcomponent scores" (p. 148). The results showed a moderate positive correlation between WM and the overall language aptitude score, and confirmed the first hypothesis of the study. However, contrary to their assumption that WM would correlate with all of the tested components of language aptitude, the findings indicated a relation only to grammatical inferencing, but not to vocabulary learning and sound-letter correspondence (Yalçın et al., 2016).

The above described studies on WM and language aptitude yielded different results regarding the association between WM and the aptitude for word learning. Therefore, the present study also addresses this issue and seeks to provide further evidence either for or against the assumption that WM correlates with an aptitude for vocabulary acquisition. As already mentioned, WM is claimed to play an important role for vocabulary learning. Hence, H1 hypothesises that people with a high working memory capacity have a higher aptitude for vocabulary acquisition. Furthermore, this study also investigates if there is any difference between students of a language and students of other fields of study. It is argued that people who are constantly engaged with languages, particularly second or foreign languages, learn new words more easily than others. As a consequence, language students are expected to be better at rehearsing items, which is crucial for the storage in the phonological store. Against this backdrop it is suggested that language students also have a higher working memory capacity. Therefore, H2 assumes that language students perform better in the aptitude test for vocabulary learning, and H3 hypothesises that language students have a greater working memory span than non-language students. Finally, as language students might have a higher working memory capacity, which is important for the development of stable representations in long-term memory, H4 expects that language students are able to recall more words in the second aptitude test for vocabulary learning than students of another field of study.

2 Methodology

2.1 Participants

All 20 participants were students. Depending on their field of study they were divided into two groups: One group comprised only language students while the other one included students of various fields of study, such as law, engineering sciences, agriculture sciences and environmental management. Each group consisted of ten subjects (50% male, 50% female). As shown in Table 1 below, the mean age of the language student group was 23.30, of the non-language student group 22.20. The range of known languages was between two and five, with the mean being 3.10 in both groups. The highest degree of education of the majority of the participants (n = 18) was the A-levels. Two subjects had already finished their master's degree, both of them were in the non-language-group. In general, the groups were

Table 1 Descriptive statistics of the participants

		Minimum	Maximum	Mean	Standard deviation	Sample size
Age	Language students	21	27	23.30	1.636	10
	Non-language students	19	26	22.20	2.150	10
Number of languages	Language students	2	4	3.10	.568	10
	Non-language students	2	5	3.10	.876	10

homogenous with respect to age, gender, known languages and highest degree of school education.

2.2 Instruments

First, participants had to complete a basic questionnaire asking for their age, gender, highest degree of school education, field of study and known languages.

Afterwards, the participants needed to take part in Part V of the Modern Language Aptitude Test (MLAT). The test, which was first published in 1959 by Carroll and Sapon, measures an individual's aptitude for learning a foreign language (Language Learning and Testing Foundation, 2014). It can be used to predict success in learning all basic communication skills, in particular speaking and listening (Language Learning and Testing Foundation, 2014). The MLAT consists of five different parts, each measuring specific skills related to foreign language learning. As the present study was focusing on foreign word learning, participants only had to complete the fifth part of the test, called Paired Associates, meaning that they had to learn a set of words from another language and memorise the English meanings of the words (Language Learning and Testing Foundation, 2014). Because all participants were native Germans with different levels of English proficiency, the test was translated into German. Participants had 2 min to memorise 24 pseudo-Kurdish words and their German meaning. Afterwards they received a list of all studied pseudo-Kurdish words and five choices in German for each of them. The participants were asked to tick the right German meaning for each pseudo-Kurdish word. There was no time limit to complete this task. To evaluate the test, all correct answers were added.

After the working memory span test, participants were tested on the 24 words of the MLAT Part V again, but without being able to look at the vocabulary list before. In this way it was tested how many words participants were able to remember after some time. The time span between the first and the second MLAT was between 10 and 20 min, depending on the time participants needed to complete the working memory test.

To test the working memory span, the Wechsler Digit Span test, which is a subtest of the revised version of the Wechsler Adult Intelligence Scale (Wechsler, 1939), and the non-word test (Benner, 2005) were used. For the former, participants heard a string of numbers which they had to repeat in the correct order, first forward and then backward. The amount of items in the test increases, from three to nine numbers for forward repetition and two to eight numbers for backward repetition. For the latter, participants listened to a string of monosyllabic non-words with an increasing number from two to eight, which they had to recall in the correct order. To score the working memory span, all accurate repetitions from both tests were counted.

2.3 Procedures

Each participant was tested on a different day either at their home or at the local music rehearsal room. The testing situation lasted between 30 and 60 min, depending on the individual's pace. After a short introduction to the study, participants first completed the questionnaire asking for personal data. After that they had to complete Part V (vocabulary learning) of the MLAT (Carroll & Sapon, 2002). Subjects were then tested on their working memory span, and finally they were asked to recall the vocabulary items of MLAT Part V again.

3 Results

To test the first hypothesis stating that people with a high working memory capacity have a higher aptitude for vocabulary acquisition, a Pearson Correlation was used. As expected, working memory span scores correlated with the first MLAT scores (Pearson's $r = 0.569$, $p = 0.009$). Unsurprisingly, the first and second MLAT scores showed a high correlation (Pearson's $r = 0.756$, $p < 0.001$). Scores of the working memory span test correlated also with the second MLAT scores (Pearson's $r = 0.466$, $p = 0.038$).

The second hypothesis assumed that language students perform better in the aptitude test for vocabulary learning. To address H2, the MLAT results of the language student group were compared with the MLAT results of the non-language student group with an independent-samples t-test. As demonstrated in Table 2 below, the scores of the first MLAT were slightly higher in the language student group ($M = 19.8$, $SD = 3.994$) than in the non-language student group ($M = 18.0$, $SD = 2.404$), however these results are not significant ($t(18) = 1.221$, $p = 0.238$).

H3, which hypothesised that language student have a greater working memory span than non-language students, was also tested with an independent samples t-test. As evident in Table 2, scores of the working memory span test of the language student-group also yielded a slightly higher result ($M = 22.3$, $SD = 4.739$) than those

Table 2 Mean scores of the first and second MLAT and the working memory test distinguished between language and non-language students

		Mean	Standard deviation	Standard error mean
MLAT V/1	Language students	19.80	3.994	1.263
	Non-language students	18.00	2.404	.760
Working memory	Language students	22.30	4.739	1.499
	Non-language students	18.90	3.107	.983
MLAT V/2	Language students	17.10	4.458	1.410
	Non-language students	16.30	3.057	.967

Table 3 Independent-samples T-test

	t	df	Sig. (2-tailed)	Mean difference	Std. error difference
MLAT V/1	1.221	18	.238	1.80	1.474
Working memory	1.897	18	.074	3.40	1.792
MLAT V/2	.468	18	.645	.80	1.709

of the non-language student-group (M = 18.9, SD = 3.107). These results are not significant (t (18) =1.897, p = 0.074).

H4 expected that language students are able to recall more words in the second aptitude test for vocabulary learning than students of another field of study, which was also tested with an independent-samples t-test. As depicted in Table 2, the second MLAT results showed that language students scored moderately higher (M = 17.1, SD = 4.458) than non-language students (M = 16.3, SD = 3.057). Again, these results are not significant (t (18) =0.468, p = 0.645).

To summarise, language students scored slightly higher in all three tests, however, the independent-samples t-tests yielded no significant results. Nevertheless, as shown in Table 3 there is a moderate trend towards language students scoring higher in the working memory span test noticeable (p = 0.074).

4 Discussion

The purpose of the current study was to provide further evidence for the assumption that WM is related to an aptitude for vocabulary learning, and to investigate whether there are differences between language and non-language students with respect to aptitude for vocabulary learning, working memory capacity, and vocabulary recall. Four hypotheses were developed to address these issues. As the studies by Yoshimura (2001) and Yalçın et al. (2016) yielded divergent results concerning the association between WM and aptitude for vocabulary acquisition, the first hypothesis was whether a high working memory capacity is associated with a high aptitude for vocabulary acquisition. The second hypothesis assumed that language students perform better in the aptitude test for vocabulary learning, and the third hypothesis expected that language students have a greater working memory span than non-language students. The fourth hypothesis assumed that language students are able to recall more words in the second attempt of the aptitude test for vocabulary learning than students of another field of study.

The results obtained support the first hypothesis, meaning that a high working memory capacity is associated with an aptitude for vocabulary acquisition. This finding corroborates the results of the study by Yoshimura (2001), which showed a significant correlation between L1 and L2 working memory scores and overall language aptitude scores, as well as the scores of the word associates and language analysis tasks. However, the results of the present study also oppose the findings of the study of Yalçın et al. (2016). Although their results indicated a moderate correlation between WM and overall language aptitude scores, they did not reveal a relationship between WM and an aptitude for vocabulary learning (Yalçın et al., 2016). Seventy-two students between the ages of 20 and 23 from an English-medium university in Turkey participated in the study of Yalçın et al. Whereas the present study used the MLAT Part V for determining the subjects' aptitude for vocabulary acquisition, the group of Yalçın et al. made use of the LLAMA, which is a computerised aptitude test by Paul Meara (2005). Because they did not only focus on the aptitude for vocabulary learning, participants had to complete all four parts of the test, namely the LLAMA B, which tests vocabulary learning, the LLAMA D, which tests sound recognition, the LLAMA E, which is a test of sound-symbol correspondence and the LLAMA F, which tests grammatical inferencing. To test working memory capacity, subjects of the present study had to recall numbers in German (L1), forwards and backwards, as well as a string of monosyllabic non-words. In contrast, the group of Yalçın et al. used two computerised reading span tasks, one in Turkish (L1) and one in English (L2), and an operation span task. They suggest that the high proficiency level of the participants might be a reason for the weak relationship between WM and language aptitude and refer to Hummel (2009) who argues that "other factors such as the amount or type of L2 exposure or motivation may be more closely related to aptitude for advanced L2 learners" (as cited in Yalçın et al., 2016, p. 153). However, as the results of the present study with participants of a similar age indicate a positive correlation between WM and aptitude for vocabulary learning, further research in this field is needed.

Regarding the comparison between the language and the non-language student-group, the study yielded no statistically significant results. Therefore, H2, H3, and H4 failed to be supported. Nevertheless, a trend exists towards language students performing slightly better in all tests. Due to the small group size and this existing trend it can be argued that it would be premature to discard H2, H3, and H4. It seems to be worthwhile to continue research in that field with bigger sample sizes and more data.

As already mentioned in the introduction, the verbal or subvocal rehearsal of items serves to refresh traces and thereby prevents them from decaying. In other words, the rehearsal system helps to maintain information and is therefore crucial for vocabulary learning. In the present study, it was expected that language students would score better in the working memory span test as well as in the aptitude test, as they are better at rehearsing because of their constant engagement with language learning. However, as mentioned above this assumption could not be supported. Nevertheless, it would have been interesting to know whether the subjects made use of rehearsing or different strategies to learn the Kurdish words and how these strategies affected the MLAT V score. The theoretical concept of the phonological loop, described in the first section, already demonstrates the importance of rehearsal in L2 vocabulary acquisition. A study conducted by Dahlen and Caldwell-Harris (2013) further supports the power of rehearsing. They addressed the question of which method of rehearsing is the most effective one. They hypothesised that the manner of rehearsal of foreign lexical items influences the subsequent recall and the recognition of these words (Dahlen & Caldwell-Harris, 2013). Participants of their study were shown Turkish vocabulary and were then required to rehearse these words according to their assigned rehearsal condition. One group was the "vocal plus auditory feedback group". Subjects of this group spoke the words aloud during the rehearsal, so they were able to hear their own voice while repeating the Turkish word. Another group was the "vocal with white noise group". Participants of this group heard white noise through earphones during the rehearsal. Consequently, participants of this group were not able to perceive the accuracy of their pronunciation and did not receive any auditory feedback. Participants of a third group, the "subvocal rehearsal group", were instructed to repeat words subvocally in their heads. Therefore, subjects of the third group did not receive auditory feedback; however, they "heard" the rehearsed words in their heads, as part of an inner speech. The fourth group was the "articulatory suppression group". Participants of this group were asked to read a few short stimulating magazine articles aloud during the presentation of the Turkish words. Consequently, subjects of the fourth group were prevented from rehearsing (Dahlen & Caldwell-Harris, 2013). The results of the study by Dahlen and Caldwell-Harris (2013) showed that participants who were able to rehearse the Turkish vocabulary were better in recalling and recognising than participants who were prevented from rehearsing. In addition, it was shown that the subvocal group performed better than the white noise group. This result suggests the important role of undisturbed rehearsal when recalling new words of a foreign language. Furthermore, the results demonstrated that the subvocal group recalled and recognised Turkish words better than or equally good as the vocal group (Dahlen & Caldwell-Harris, 2013). Dahlen and Caldwell-Harris explain the powerful effect of subvocal rehearsal by stating that silent rehearsal in inner speech is faster than vocal articulation, allowing the test words to be rehearsed more often in the phonological loop than it is possible with slower vocal rehearsal. They also suggest that overt vocalisation may actually detract from learning, as the subject has to divide his or her attention between the inner processing function of rehearsal and the performance aspect of overt pronunciation. They argue that this might explain why children who are forced to read aloud in class may understand little of the content of their reading (Dahlen & Caldwell-Harris, 2013).

5 Conclusion

Vocabulary acquisition is subject to individual differences. Two variables that influence the vocabulary development are WM capacity and aptitude for word learning. The results of the present study indicate a correlation between these two variables. Nevertheless, divergent findings of various studies in this field show that further research is needed to examine the association between WM and aptitude for vocabulary acquisition.

As it is the case in every study, this one had its limitations. For instance, the small sample size is a limitation in regard to group comparisons, for which usually more participants are recommended. Another limitation is the missing information about the way of how subjects studied the Kurdish words. Answers to this unasked question could provide interesting information about whether participants made use of rehearsing or other vocabulary learning strategies. By knowing how many language and non-language students made use of rehearsing, it might have been possible to support the proposed assumption that language students are better at rehearsing. A last limitation of the study is the time span between the two aptitude tests, which was probably too short to yield significant results.

Despite these limitations, the current study supports the hypothesis that WM affects aptitude for vocabulary acquisition. Future research should replicate H2, H3, and H4 with a greater sample size in order to further investigate differences between language and non-language students. Additionally, it would be helpful to investigate whether there is a difference in learning strategies between language and non-language students to disclose cognitive differences between these groups.

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, MI: Erlbaum.

Baddeley, A. (1986). *Working memory*. London: Oxford University Press.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2010). Working memory. *Current Biology, 20*(4), 136–140. https://doi.org/10.1016/j.cub.2009.12.014

Baddeley, A., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158–173.

Benner, U. (2005). *Syllables in speech production – A study of the mental syllabary*. Linguistic Master thesis at the University of Stuttgart, Stuttgart.

Carroll, J. B., & Sapon, S. (2002). *Modern language aptitude test*. Bethesda, MD: Second Language Testing.

Chrysochoou, E., Bablekou, Z., Masoura, E., & Tsigilis, N. (2013). Working memory and vocabulary development in Greek preschool and primary school children. *European Journal of Developmental Psychology, 10*(4), 417–432. https://doi.org/10.1080/17405629.2012.686656

Cowan, N. (1997). The development of working memory. In N. Cowan (Ed.), *The development of memory in childhood* (pp. 163–199). London: Psychology Press.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press.

Dahlen, K., & Caldwell-Harris, C. (2013). Rehearsal and aptitude in foreign vocabulary learning. *The Modern Language Journal, 97*(4), 902–916. https://doi.org/10.1111/j.1540-4781.2013.12045.x

Dewaele, J. (2013). Learner-internal psychological factors. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge handbook of second language acquisition* (pp. 159–179). Cambridge, UK: Cambridge University Press.

Dörnyei, Z. (2005). *The psychology of the language learner: individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*, 42–68. https://doi.org/10.1075/aila.19.05dor

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell.

Gathercole, S. E., & Baddeley, A. D. (1993). *Working memory and language*. Hove, UK: Lawrence Erlbaum.

Hummel, K. (2009). Aptitude, phonological memory, and second language proficiency in novice adult learners. *Applied PsychoLinguistics, 30*, 225–249. https://doi.org/10.1017/S0142716409090109

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching, 44*, 137–166. https://doi.org/10.1017/S0261444810000509

Language Learning and Testing Foundation. (2014). http://lltf.net/aptitude-tests/language-aptitude-tests/modern-language-aptitude-test-2/. Accessed 5 Feb 2016.

Meara, P. (2005). *LLAMA language aptitude tests*. Swansea, UK: Lognostics.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge, UK: Cambridge University Press.

Morra, S., & Camba, R. (2009). Vocabulary learning in primary school children: Working memory and long-term memory components. *Journal of Experimental Child Psychology, 104*, 156–178. https://doi.org/10.1016/j.jecp.2009.03.007

Pickering, S., & Gathercole, S. E. (2001). *Working memory test battery for children (WMTB-C)*. London: The Psychological Corporation.

Repovs, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience, 139*, 5–21. https://doi.org/10.1016/j.neuroscience.2005.12.061

Robinson, P. (2002). Learning conditions, aptitudes complexes, and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 113–136). Philadelphia: John Benjamins.

Rota, G., & Reiterer, S. M. (2009). Cognitive aspects of pronunciation talent. In G. Dogil & S. M. Reiterer (Eds.), *Language talent and brain activity* (pp. 67–96). Berlin, Germany/New York: Mouton de Gruyter.

Speidel, G. E. (1989). A biological basis for individual differences in learning to speak. In G. E. Speidel & K. E. Nelson (Eds.), *The many faces of imitation in language learning* (pp. 199–229). New York: Springer.

Towell, R. (2013). Learning mechanisms and automatization. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge handbook of second language acquisition* (pp. 114–136). Cambridge, UK: Cambridge University Press.

Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology, 141*, 65–82. https://doi.org/10.1016/j.jecp.2015.06.015

Wechsler, D. (1939). *The measurement of adult intelligence* (p. 229). Baltimore: Williams & Wilkins.

Wen, Z. (2012). Foreign language aptitude. *ELT Journal, 66*(2), 233–235. https://doi.org/10.1093/elt/ccr068

Yalçın, S., Çeçen, S., & Erçetin, G. (2016). The relationship between aptitude and working memory: an instructed SLA context. *Language Awareness, 25*(1–2), 144–158. https://doi.org/10.1080/09658416.2015.1122026

Yoshimura, Y. (2001). The role of working memory in language aptitude. In X. Bonch-Bruevich, W. J. Crawford, J. Hellermann, C. Higgins, & H. Nguyen (Eds.), *The past, present, and future of second language research* (pp. 144–163). Sommerville, MA: Cascadilla Press.

Working Memory and Language Aptitude with Focus on L2 Vocabulary Learning

Hyun Jung Kim

Following approaches of the precedent researches that examined the correlation between measures of working memory (WM) test and classical language aptitude test such as MLAT, this study has a main objective to probe compatibility of WM capacity task to language aptitude test with focus on L2 vocabulary learning. It was hypothesized that a correlation could be found between results of WM capacity task and chosen subtests of language aptitude test which are mainly engaged in vocabulary learning. Considering the fact that it is one of major findings in previous WM studies that overall WM capacity of children is lower than adults, another incidental research question was included whether it can be also verified by WM test results of this study. 18 Korean L2 English learners (users) were selected to take the following three sets of test: (1) Combined WM test of Digit Span task and Backward Digit Span task; (2) Language aptitude test – MLAT V; (3) Language aptitude test – LLAMA B. The results showed that WM capacity was significantly correlated with test score of MLAT V, whereas it was not correlated with LLAMA B. Further, no significant difference between the two groups was shown in WM tests, while t-test result leaves a certain expectation that a more feasible outcome can be derived with an increased number of sample.

1 Introduction

Memory has always been an intriguing theme when it comes to vocabulary learning in the field of SLA research because "unlike grammar, which is a system of a limited number of rules, vocabulary is an open set of many thousands of items" (Laufer

H. J. Kim (✉)
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria
e-mail: a01277000@unet.univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_4

& Nation, 2012, p. 163), and one should acquire each item one by one to reach a certain level of L2 lexical inventory by active utilization of his/her memory. Thus one's proficiency in second language vocabulary learning is often regarded as his/her efficiency in acquiring novel words in a limited period of time as well as retaining them for a longer period in his/her long-term memory as a reservoir. Like in any other domains of second language learning individual differences are also shown in vocabulary learning, and this means that one can pick up certain amount of words and hold them in his/her memory more proficiently than others (Morra & Camba, 2009). In the field of SLA research it has long been a subject of study why some learners show better performance in their second language learning than others, and under this background conceptualization of language aptitude has been borne. A motivation behind this study goes just in line with this, and its main concern is to explain varying degree in individual proficiencies in L2 vocabulary learning with regard to language aptitude in the relevant framework of memory.

As well known, Carroll (1981, p. 105) presented four-factor language aptitude theory which includes the following four components:

(i) Phonemic coding ability – the ability to identify distinct sounds, to form associations between the sounds and symbols representing them, and to retain these associations;
(ii) Inductive language learning ability – the ability to infer or induce the rules governing a set of language materials, given sample language materials that permit such inferences;
(iii) Grammatical sensitivity – the ability to recognize the grammatical functions of words or other linguistic entities in sentence structures; and
(iv) Rote learning ability, namely associative learning – the ability to learn associations between sounds and meanings rapidly and efficiently, and to retain these associations

These are all reflected on MLAT, an influential language aptitude test elaborated by himself with his colleague (Carroll & Sapon, 1959). The sort of memory engaged here as a focus is associative memory, the ability to learn and remember the relationship between unrelated items, which was the form of memory dominant in psychology at the time of development of the test (Skehan, 2012, p. 384). This conceptualization of memory was later criticized as being somehow limited, however, in that "the capacity to analyze larger quantities of material rather than limited quantities that can be held by short-term memory, and the capacity to memorize material lacking in familiarity as opposed to a concept of utilizing associative memory are also relevant to aptitude" (Skehan, 2012, p. 384).

Significant development has been made in the field of memory since the age of Carroll's conceptualization of language aptitude and with much wider interpretations of memory at hand, what stands at the forefront of memory studies now is WM (Skehan, 2012, p. 384). First proposed by Baddeley and Hitch in 1974, WM has been one of the most innovative research themes in brain science in the last decades and "its construct has been the focus of extensive research in cognitive psychology and psycholinguistics" (Juffs & Harrington, 2011, p. 137; Baddeley, 2007;

Gathercole, Pickering, Ambridge & Wearing, 2004; Unsworth, Hertz & Engle, 2005). Its relevance and centrality in language aptitude was strongly suggested by Miyake and Friedman (1998) with their "WM as language aptitude" hypothesis, and based on this, Skehan later presented a conjecture of various stages of L2 processing each of which is linked with potential WM involvement as follows (Skehan, 2012, p. 385):

(i) Input processing (segmentation) – more phonological memory enables longer stretches of language to be processed, and parsing therefore to be more efficient;
(ii) Noticing and handling form and meaning simultaneously – greater capacity can enable parts of input to be extracted;
(iii) Pattern identification (recognition) – more input available enables patterns of greater length to be identified;
(iv) Complexification/Restructuring (pattern restructuring and manipulation) – more capacity enables connections to be made between current WM and what is held in long-term memory, as well as to enable long-term memory to be changed;
(v) Error avoidance – more WM capacity enables attention to be directed to monitoring and error avoided;
(vi) Response to feedback – more memory enables attention to be directed to feedback, and the incorporation of feedback into performance, as well as the potential to change long-term memory; and
(vii) Automatization/Lexicalization – more material in WM enables chunking which can be transferred to long-term memory.

WM is generally referred to as "system that is used for the temporary maintenance of task-relevant information whilst performing cognitive tasks" (Williams, 2012, p. 427). Despite there still being different WM models competing each other, a general consensus has been reached in the field of cognitive psychology with regard to conceptualization of its construct related to language learning (Wen & Skehan, 2011, pp. 21–22; Gathercole & Baddeley, 1993; Kellogg, Olive & Piolat, 2007; Majerus, Poncelet, Greffe, & Van der Linden, 2006; Majerus, Poncelet, Van der Linden & Weekes, 2008; Papagno & Vallar, 1995). Most fundamental part of this consensus is related to its structure. WM is a multi-component cognitive construct with two constituents of (i) slave system that works as a short-term storage, comprised of phonological loop and visuo-spatial sketchpad, each of which handles phonological/verbal and visual/spatial information respectively; (ii) central executive that controls and coordinates the flow of information between the two elements of slave system above. Later episodic buffer was added to the slave system as another element which works as a place where the various types of information are temporarily stored and integrated to form episodes by using multi-dimensional coding and communicates with long-term memory.

With regard to different components of WM introduced above, there are different research groups with varying foci of research. For example, cognitive psychologists in Britain such as Baddeley and Gathercole have mainly focused their attention to the phonological loop component and investigated its involvement in different aspects of language learning, vocabulary acquisition in particular (Wen & Skehan, 2011, p. 22). There has also been another research group in North America, on the other hand, members of which have generally placed more emphasis on the central executive component. Represented by scholars like Daneman and Miyake, this group concerned the individual differences in the central executive function and its effects on language processing and comprehension (Wen & Skehan, 2011, p. 24).

The next concern may well be then how one can measure WM capacity. In principle both storage and processing components of WM can be measured separately or in combination. The type of tests that deal with the former is usually called simple WM test. A simple short-term memory capacity is typically "measured by the number – or span – of unrelated digits or words that can be recalled" (Juffs & Harrington, 2011, p. 141). Considering the fact that both digit and word span tests require some knowledge of a certain language, an alternative method of Non-Word Repetition task is also used to lessen this previous language knowledge effect (Juffs & Harrington, 2011, p. 141). Variables to affect Non-Word Repetition performance are abilities to "perceive, store, recall and reproduce phonological sequences" (Juffs & Harrington, 2011, p. 141), and a majority of researches that worked on the correlation between vocabulary learning and working memory either in L1 or L2 used this type of tests (Williams, 2012, pp. 428–429). On the other hand, there are tests to measure both storage and processing component – phonological memory and central executive – together. Backward Digit Span task is categorized into this category, and a possible task here, for example, is to recall numbers of 2 to 8 digits backwards. It is not a mere storage of digits in memory, but a processing of them by having numbers backwards in mind to recall previous numbers while retaining recent ones. This type of test is referred to as complex WM test, and it includes Reading Span Test that "assesses individual's ability to simultaneously read and comprehend a set of sentences and then recall a target word for each, usually the last word in the sentence" (Juffs & Harrington, 2011, p. 142). Another type of complex WM test, Listening Span Test follows the same format but requires test-takers to listen to the sentences instead of reading.

With regard to the correlation between WM capacity and vocabulary learning, which constitutes main concerns of this study, a series of studies have proven that phonological short-term memory as a component of WM plays an important role in L1 and L2 learning, especially in the development of vocabulary learning. In the field of SLA research in specific, it has been proved that the role which phonological memory plays in learning new sound patterns is critical to L2 vocabulary learning (Juffs & Harrington, 2011, p. 140). Among them is a research that revealed a link between Non-Word Repetition test score and test performance on the L2 English vocabulary learning with Finnish schoolchildren (Service & Kohonen, 1995). Another research showed similar results with L1 Chinese high school students of English vocabulary learning but only among the lower proficiency group (Cheung, 1996). Speciale, Ellis and Bywater (2004) also proved the importance of phonological short-term memory capacity by linking it to the ability to learn novel L2 Spanish vocabulary of L1 English university students. Kormos and Sáfár (2008) also examined the correlation between success of acquisition of various language skills and competencies during an intensive English language training program held for one year and WM capacity measured from both simple phonological memory capacity task (Non-Word Repetition task) and complex WM task (Backward Digit Span task). The result showed that there was no correlation between the success of second language acquisition and simple WM task, while there was between Backward Digit Span task.

While these studies contemplating on the correlation between phonological short-term memory or simple/complex WM and L2 vocabulary learning mainly concerned output or performance of learning, often operationalized as scores in proficiency test, there has been a group of studies that examined the relationship between different aspects of memory and language aptitude itself. Approach of this study is just in line with this type of test model. In these studies a correlation was examined between measures of WM test and the subsets of classical language aptitude test such as MLAT. Hummel (2009), for instance, found no significant correlation between phonological short-term memory operationalized by Non-word Repetition task and subcomponents of MLAT, despite the fact that each test was found to be predictive on overall learning results. Robinson (2002) presented a contrasting result, however, when he examined the correlation between complex Reading Span task and subtests of MLAT. These studies adopted MLAT test in common as language aptitude measures, but differed from each other in that one took simple WM task while the other complex. Following these earlier studies, this study also chose a subset of MLAT test as one of two language aptitude tests, while it administrated complex type of WM task.

This study is also interested in the difference between children and adults of their WM capacity, and with regard to the three components of original WM model of Baddeley's, corresponding to phonological loop, visuo-spatial sketchpad and central executive, there was an inspiring study that probed a question if the 'final state' of WM system in adults might differ from 'developing' one in children (Gathercole, Pickering, Ambridge, & Wearing, 2004). It was postulated that children might have a more nonspecialized WM system than adults, and the division of three components might be rather blurred in children. The study assessed three measures from each WM component of over 700 children aged 4 to 15 years to compare closeness among measures of each component with that among measures of other components. Its result showed that the former is stronger than the latter, to prove that the structure of WM system of children and adults are the same. On the way of drawing this conclusion one interesting result was also found that all the measures of each component presented substantial increases with age, reaffirming a large amount of previous literature (Henry, 2012, p. 127).

Evidently, it has been emphasized in many precedent researches that WM can function as a key component in foreign language aptitude with its robust role, and there has been a controversy also on how far WM capacity can be regarded as a trait, "a stable individual characteristic that ultimately has a neurocognitive basis" as opposed to a dynamic state, "whose capacity and efficiency is affected by factors such as domain knowledge and learner goals" (Juffs & Harrington, 2011, p. 156).

Originally, a group of researches that examined relationship between WM capacity and language aptitude stands on the position that WM is a rather stable individual feature over time. There is an opposing side of view, however, that suggests WM is something modifiable and can be improved through experience and training. They argue that the scope of capacity modification can be stretched from phonological and visual memory to executive function.

Following the approaches of the precedent researches above, this paper, main concern of which is the correlation between WM capacity and language aptitude in one's vocabulary learning, will examine the correlation between WM capacity measurements and two different language aptitude tests, which are regarded as mainly engaged in vocabulary learning. Additionally, it will address an incidental issue whether WM test results show a meaningful difference between the subgroups of children (school-aged to adolescents) and adults to verify a general finding of precedent WM researches that WM capacity of children is lower than that of adults. Two research questions of this study can be put as follows:

1. Is WM capacity correlated with language aptitude in ones' L2 vocabulary learning?
2. Is WM capacity of children (school-aged to adolescents) lower than adults?

2 Methodology

2.1 Participants

In this study 18 Korean L2 English learners (users) were selected, eight of whom were male and ten female (aged 10–45; mean = 28.11; SD = 14.55). All participants were native Korean speakers who had learned English as a second language with varying age of onset. (most adults at the age of 13; most children at 5–6). Two of the participants lived in monolingual setting while the rest in bilingual/trilingual (English and/or German at work/school and Korean at home). To compose a set of two subgroups (school-aged children and adolescents/adults) more efficiently, recruitment was mostly made on a family basis to have parents and their children as participants of the test at the same time. The children group (n = 9) includes school-aged children and adolescents aged from 10 to 18, while the participants of the adults group (n = 9) were aged from 38 to 45. For the breakdown of the characteristics of the two groups, please refer to the table below (Table 1).

All of the participants completed the following tests: WM tests including digit span forward and digit span backward test, language aptitude test MLAT V, and language aptitude test LLAMA B.

Table 1 Characteristics of participants

	Children (n = 9)		Adults (n = 9)	
Age	School-aged	Adolescents	18–29	30–49
	4	5	0	9
Gender	Male	Female	Male	Female
	5	4	3	6
Lingual setting	Monolingual	Bi-/trilingual	Monolingual	Bi-/trilingual
	1	8	2	7
Age of onset	6–12	13–19	6–12	13–19
	9	0	1	8

2.2 Procedures

All participants tested individually over a period of 1–2 days in a quiet room at home, so that each individual could take the test in a manner that he/she responds immediately and operates at optimum capacity, considering the fact that accurate measurement of WM capacity is often confronted by considerable challenges with regard to ways how tests are administered (Juffs & Harrington, 2011, p. 143). All the adult participants took all the tests in one day, while some younger child participants in 2 days due to lack of concentration. Four tests were administered to each participant: two WM tests (digit span forward recall and digit span backward recall) and two language aptitude tests (MLAT V and LLAMA B). The tests were administered in a set sequence and the order of test was held constant (digit span forward test – digit span backward test – MLAT V – LLAMA B), considering the nature of memory demands of each test and its mode of testing (paper-based or computer-based). One WM test involved verbal storage only and is associated with the phonological loop among the three WM components of Baddeley's (digit span recall test), while the other WM test, digit span backward recall test, involved complex memory, being associated with both phonological loop and central executive. One language aptitude test involved associative memory between unfamiliar word and its meaning (MLAT V), while the other involved unfamiliar object (visual image) and its name.

In order to test WM capacity of the participants this study used a WM test composed of both digit span forward and digit span backward recall for which the participants are requested to recall strings of numbers from 3 to 9 (forward) and 2 to 8 digits (backward) from his/her memory in correct serial order. A list of digit sequences (14 sequences for each span forward and backward) was constructed randomly and each sequence was read aloud to the participants at a rate of one digit per second. Numbers were read only one time, and each participant's recall of each string was promptly evaluated and scored by the test administrator. If a participant made a mistake, an additional number string with the same length was read to him/her. If second trial succeeded, he/she can move on to the next strings which were one digit longer), while the session ended when both trials failed.

For the measurement of language aptitude of the participants this study used two types of subtest from extensive language aptitude tests, which are MLAT V

(Carroll & Sapon, 1959) and LLAMA B (Meara, 2005). Developed as omnibus aptitude test batteries already in 1950s, MLAT has been widely used in various researches and recognized as one of the most influential and comprehensive tests for assessing one's language aptitude. While its five subsets, Number Learning (MLAT I), Phonetic Script (MLAT II), Hidden Words (MLAT III), Words in Sentences (MLAT IV) and Paired Associates (MLAT V) are not always direct measures of the four language aptitude constructs that Carroll conceptualized (Skehan, 2012, p. 382), MLAT V which is to test rote learning ability of individual by making links between two verbal elements (English and Kurdish words) well matches one, corresponding to associate learning. The participants were first given printed sheets showing a total of 24 word pairs to memorize and told to find the right English counterpart of each Kurdish word after 2 min of learning.

Another subset of aptitude test used in the study was LLAMA B, which is a relatively new aptitude measure and is differentiated from MLAT in that it is computer-based and language-independent by using picture stimuli (Granena, 2011; Rogers, Meara, Barnett-Legh, Curry & Davie, 2017). Loosely based on MLAT, the test includes four subtests which measure paired associates vocabulary learning (LLAMA B), phonemic memory capacity of unfamiliar sound sequences (LLAMA D), ability for detecting sound-symbol correspondence (LLAMA E) and inductive grammatical learning (LLAMA F). In this study the participants took LLAMA B where the names of 20 visual images of unusual objects were shown simultaneously on screen for 2 min with name of each to learn (the name of an object is displayed upon clicking on it). The program places no constraints on how to do this, so participants can adopt any different strategies to finalize the mission. At the end of the learning phase, participants are asked to pick the right name for each object being identified by clicking on it.

3 Results

3.1 Correlation: WM Capacity and Language Aptitude

To answer the first research question of this study, a correlation between the WM measurements and language aptitude figures of the participants was tested, and it was found that WM capacity was significantly correlated with MLAT V, $r = 0.56$, p (two-tailed) < 0.05. However, WM capacity was not correlated with another language aptitude test, LLAMA B (p (two-tailed) > 0.05) (Table 2).

Table 2 Correlations (Pearson)

	Working memory	MLAT V	LLAMA B
Working memory	1	.564*	−.106
MLAT V	.564*	1	.239
LLAMA B	−.106	.239	1

*Correlation is significant at the 0.05 level (2-tailed)

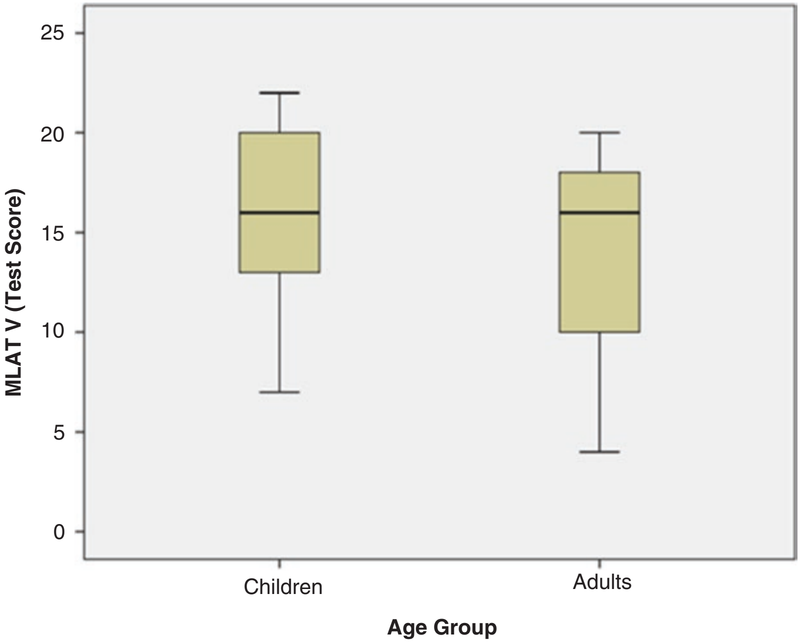


Fig. 1 Language aptitude test: MLAT V

3.2 Language Aptitude Test: MLAT V and LLAMA B

Even though there was no research question related to language aptitude tests alone for this test, a glimpse on the results reveals the following figures. The mean of the participants' scores of MLAT V was 15.22, SD = 4.78, and the scores ranged from 4 to 22 (maximal score 24). The mean of the children group was slightly higher with 16.00 (SD = 5.03, score range from 7 to 22) than that of adults group with 14.44 (SD = 5.39, score range from 4 to 20). The mean of the LLAMA B scores of the participants was calculated to be 7.17 (SD = 5.12), with scores ranging from 1 to 16 (maximal score 20). The children group showed the higher mean (mean = 8.78; SD = 5.40) as in the case of MLAT V with scores ranging from 2 to 16, as compared to the adults group (mean = 5.56; SD = 3.68) with scores ranging from 1 to 11 (Figs. 1 and 2).

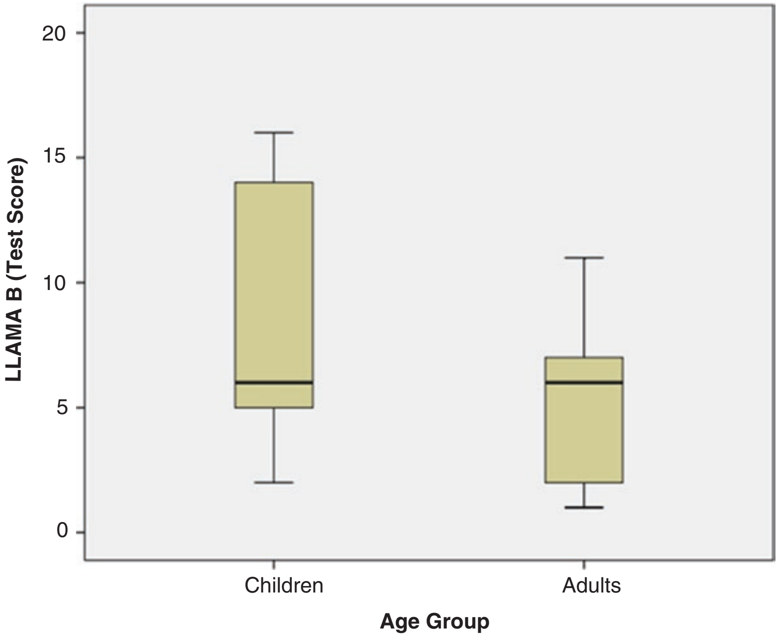


Fig. 2 Language aptitude test: LLAMA B

3.3 WM Capacity: School-Aged Children and Adolescents vs. Adults

The mean of the WM capacity test scores of each participant was 18.83, SD = 6.44. The lowest score was 10 and the highest score was 28 (maximal score 28). The school-aged children and adolescents group showed the mean of 16.11 (SD = 5.40) with the highest score 25 and the lowest 10, which was much lower than the adults group's 21.56 (SD = 6.50) with the highest score 28 and the lowest 10. Except for the two outliers with the highest scores of 25, the former group had lower scores in lower to upper quartile, ranging from 13 to 16 with lower level of deviation, while latter group showed much higher scores in lower to upper quartile, ranging from 19 to 26, with the whole set of scores stretching over a much wider range from the lowest 10 to highest 28.

To answer the second research question of this paper t-test was performed, while it might not be feasible enough, given the limited number of sample for each group (n = 9 respectively). But it might present, at least, a possible clue to interpret the difference in distribution shown between the two groups. The result was that there was no significant difference between the two groups found, however, while the calculated p-value = 0.072 (>0.05) leaves a certain expectation that a more meaningful outcome can be derived with an increased number of each sample (Fig. 3).

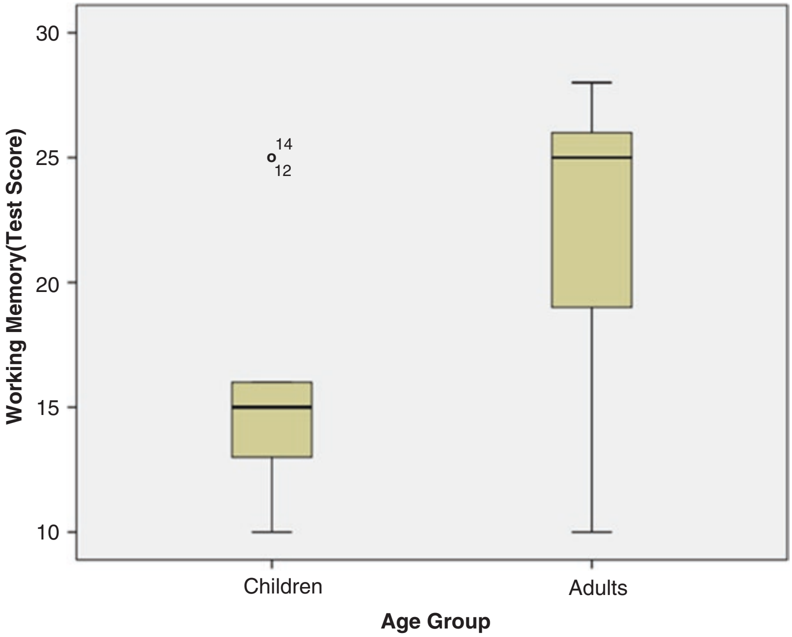


Fig. 3 Working memory capacity test

4 Discussion

With regard to its main concern to examine the correlation between WM capacity and measured language aptitude with focus on vocabulary learning, this study found that WM capacity correlates with MLAT V, taken as measurement for vocabulary learning language aptitude. Among a series of studies that examined the relationship between the two, this result confirms Robinson's (2002) who found the correlation between complex WM test (Reading Span task) and subtests of MLAT. Another language aptitude test, LLAMA B, however, was not found to be correlated with WM capacity. In the light of these results, it may be worth being discussed that we must consider that the different types/aspects of WM involve in different types of language aptitude, if we are to better understand the relationship between WM and language aptitude testing (Dörnyei & Skehan, 2003; Li, 2016; Robinson, 2005; Skehan, 2016; Wen, Biedroń, & Skehan, 2017).

A considerable list of literature clearly illustrates that WM is not a unitary construct, but rather a set of processes that involve in different stages of language learning. It performs dual functions by combining storage with processing and manipulation of information with its two sub-constituents – slave system and central executive – and importance of each function varies according to L2 domain. In this view, "the effect of WM capacity as a constraint on L2 performance will differ" in vocabulary learning domain, for instance, from that of sentence processing or text comprehension. How different WM subsystems interact in each domain, however, remains to be answered (Juffs & Harrington, 2011, p. 159). Similar to this, Wen argued that different types of WM are linked to different types of language aptitude, and proposed the "Integrated Approach" in which phonological working memory is a "language learning device" and executive working memory is involved with "language processes" (Wen 2016, p. 147). Following these hypotheses of the earlier studies, the main task of this study can be restated as follows: which WM functions affect one's language aptitude in his/her vocabulary learning? As the participants of this study were tested on their WM capacity through complex WM test for which two WM subsystems – phonological loop and central executive – involve, and their test scores were strongly correlated with their scores of vocabulary learning test of MLAT V, this can be interpreted that language aptitude which engages in vocabulary learning is affected by the two functions of WM, that are phonological memory and executive memory.

Then a remaining issue would be how we can interpret the result that there was no correlation found between the WM task and another language aptitude test, LLAMA B. It might be argued that a combination of phonological memory and executive memory that linked to MALAT V does not affect LLAMA B. One possible clue to interpret this might come from the fact that LLAMA B task includes a type of inputs that is different from MLAT V. Those are visual images of unfamiliar objects which are stored at visuo-spatial sketchpad, separated from phonological loop that stores verbal/phonological input in one's WM system. The difference between the two tests, MLAT V and LLAMA B that the former presents verbal pairs as associate learning task, while the latter visual and verbal pairs may explain the discrepancy in correlations of each language aptitude test measurement with WM scores. In this sense, it can be argued that LLAMA B links to memory functions beyond those compatible with MLAT, even though its origination was loosely based on the latter. And to better understand the very mechanism of memory functioning in LLAMA B, a further catered version of complex WM task which also includes visual memory storage and its processing should be conceived.

As research on phonological memory has been a focus of WM studies in SLA research so far, there is only a vague picture of how visual information is processed in different language learning domains. LLAMA B was a good example to reveal some implication in this regard. Among the three sub-components of the slave system in WM, visuo-spatial sketchpad and episodic buffer have received less attention compared to phonological loop from language researchers, reflecting the dominant role it plays in processing of both spoken and written language. This is partly because most of the WM capacity researches have been carried out so far in alphabet-based languages where phonological information is especially important. Still, a new wave is recently detected that the importance of visual-orthographic skills in written language processing of non-alphabetical language, Chinese in specific, is highlighted. But episodic buffer has not been even dealt with in language studies with practically no research on possible roles that this component might play in one's L2 learning and use (Williams, 2012, p. 437). By paying more attention to specific WM components, however, "more direct, non-linguistic, tests of executive function" can be elaborated to be usefully applied to further examination of the relationship between WM and SLA. For this, achieving "far more evidence of the relationship between different aspects of WM and specific learning processes" must be an urgent prerequisite (Williams, 2012, p. 438).

This might also shed lights on an issue related to the second research question of this study if there is a significant difference of WM between children and adults, that is, if one's WM system goes through developmental phases in childhood (Harley & Hart, 1997). The study that measured WM capacity functioned by its phonological loop and central executive component didn't confirm a majority of earlier studies that showed children are outperformed by adults in WM capacity, even though it revealed a slight trend of distinction between the children and adults group. Apparently, many interesting future research themes might arise when our attention is paid to different WM components/functions of children, visuo-spatial sketchpad and episodic buffer functions in specific, as discussed earlier regarding the first research question. An extended research question that adds language aptitude to this will be also valid. In language learning aptitude research, it has also been one of the fundamental assumptions behind its concept that it is a relatively stable characteristic and is not changeable by training or influenced by any experience (Skehan, 1998). There are opposing views to this, however, arguing that "language aptitude is a form of developing expertise rather than an entity fixed at birth" (Grigorenko, Sternberg & Ehrman, 2000, p. 401). A complex dynamic nature of language learning aptitude was reaffirmed recently by Dörnyei (2010) with an argument that it can be "affected by a range of internal and external learner factors, similar to motivation" (Kormos, 2011, p. 144).

5 Conclusion

It is widely believed that the concept of foreign language aptitude that refers to an individual's ability or talent in the process of language learning is still valid, and what has been yet changed is our understanding of its composing components, their inter-relationship, and most of all, incorporating WM as one of its key components (Skehan, 2012, p. 391). This paper tried to demonstrate this by reviewing the compatibility of WM capacity task to language aptitude test by examining the correlation between them with two different yet related sorts of existing language aptitude tests in specific language learning domain of vocabulary learning. A meaningful result has been derived from a small-sized empirical study that WM, working as a part of cognitive mechanism rather than a mere memory, was found to be strongly correlated with another dimension of cognitive ability of individual, that is, language aptitude (MLAT V). But to better understand the compatibility between them, and thus to conceive of a more comprehensive model of language aptitude measurement where WM is implicated as its core element, more careful attention to its construct itself is needed, as seen in the case of LLAMA B. Still, it must be noted as a critical shortcoming of this study that a standard version of tests that were originally invented for adults was presented also to the children during each test administration. Also, a limited size of sample and less comprehensive WM capacity task used for the study as well as insufficient references that dealt with validation of LLAMA are all limitations of this study and left to be resolved for more fruitful discussions.

References

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford, UK: Oxford University Press.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual difference and universals in language learning aptitude* (pp. 119–154). Rowley, MA: Newbury House.

Carroll, J. B., & Sapon, S. (1959). The Modern Languages Aptitude Test. In Z. Dörnyei (Ed.)., (2006) Individual differences in second language acquisition. *AILA Review* (Vol. 19, pp. 42–68).

Cheung, H. (1996). Non-word span as a unique predictor of second language vocabulary learning. *Developmental Psychology, 32*, 867–873.

Dörnyei, Z. (2010). The relationship between language and language learning motivation: Individual differences from a dynamic systems perspective. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247–267). London: Continuum.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell.

Gathercole, S. E., & Baddeley, A. (1993). *Working memory and language*. Hove, UK: Lawrence Erlbaum Associates.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory from 4 to 15 Years of Age. *Developmental Psychology, 40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Granena, G. (2011). Cognitive aptitudes for second language learning and the LLAMA Language Aptitude Test. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 105–130). Amsterdam: John Benjamins Publishing Company.

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A Theory-based Approach to the Measurement of Foreign Language Learning Ability: The Canal-F Theory and Test. *The Modern Language Journal*, 84(3), 391–405. https://doi.org/10.1111/0026-7902.00076

Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 19*(3), 379–400.

Henry, L. (2012). *The development of working memory in children*. Los Angeles: Sage.

Hummel, K. M. (2009). Aptitude, phonological memory, and second language proficiency in nonnovice adult learners. *Applied Psycholinguistics, 30*, 225–249. https://doi.org/10.1017/S0142716409090109

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching, 44*(2), 137–166.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica, 124*(2007), 382–397. https://doi.org/10.1016/j.actpsy.2006.02.005

Kormos, J. (2011). New conceptualizations of language aptitude in second language attainment. In G. Granena & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 131–152). Amsterdam: John Benjamins Publishing Company.

Kormos, J., & Sáfár, A. (2008). Phonological short term-memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*(2), 261–271. https://doi.org/10.1017/S1366728908003416

Laufer, B., & Nation, I. S. P. (2012). Vocabulary. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 163–176). New York: Routledge.

Li, S. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition, 38*(4), 801–842. https://doi.org/10.1017/S027226311500042X

Majerus, S., Poncelet, M., Greffe, C., & Van der Linden, M. (2006). Relations between vocabulary development and verbal short-term memory: The relative importance of short-term memory for serial order and item information. *Journal of Experimental Child Psychology, 93*(2), 95–119. https://doi.org/10.1016/j.jecp.2005.07.005

Majerus, S., Poncelet, M., Van der Linden, M., & Weekes, B. S. (2008). Lexical learning in bilingual adults: The relative importance of short-term memory for serial order and phonological knowledge. *Cognition, 107*(2), 395–419. https://doi.org/10.1016/j.cognition.2007.10.003

Meara, P. (2005). *LLAMA Language Aptitude Tests*. Swansea: Lognostics.

Miyake, A., & Friedman, D. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention*. Mahwah, NJ: Lawrence Erlbaum Associates.

Morra, S., & Camba, R. M. (2009). Vocabulary learning in primary school children: Working memory and long-term memory components. *Journal of Experimental Child Psychology, 104*(2), 156–178. https://doi.org/10.1016/j.jecp.2009.03.007

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *The Quarterly Journal of Experimental Psychology, 48*(1), 98–107.

Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on incidental second language learning: A replication and extension of Reber, Walkenfield and Hernstadt (1991). In P. Robinson (Ed.), *Individual differences and second language instruction* (pp. 211–266). Philadelphia: Benjamins.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73. https://doi.org/10.1017/S0267190505000036

Rogers, V., Meara, P., Barnett-Legh, T., Curry, C., & Davie, E. (2017). Examining the LLAMA aptitude tests. *Journal of the European Second Language Association, 1*(1), 49–60. https://doi.org/10.22599/jesla.24

Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics, 16*, 155–172.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, NY: Oxford University Press.

Skehan, P. (2012). Language aptitude. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 381–395). New York: Routledge.

Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in L2 processing and acquisition*. Amsterdam: John Benjamins.

Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*(2), 293–321. https://doi.org/10.1017/S0142716404001146

Unsworth, N., Hertz, R. P., & Engle, R. W. (2005). Working memory capacity in hot and cold cognition. In R. E. Engle, G. Sedek, U. von Hecker, & D. N. McIntosh (Eds.), *Cognitive limitations in aging and psychopathology* (pp. 19–43). New York: Cambridge University Press.

Wen, Z. (2016). *Working memory and second language learning: Towards an integrated approach*. Bristol, UK: Multilingual Matters.

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting a case for working memory as language aptitude. *A Journal of Language and Literature, 60*, 15–43. https://doi.org/10.5007/2175-8026.2011n60p015

Wen, Z. E., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching, 50*, 1), 1–1),31. https://doi.org/10.1017/S0261444816000276

Williams, J. N. (2012). Working memory and SLA. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 427–441). New York: Routledge.

Part II
Language Aptitude and Psychological Factors

On the Role of Self-Efficacy as a Possible Component of Language Aptitude in the Acquisition of British [æ]

Daniel Leisser

Abstract In social cognitive theory it is a common assumption that individuals are not merely passive recipients of external factors, but are able to form and affect their environment. Self-efficacy has been described as one of the major factors seemingly connected with the development of linguistic competence. A considerable amount of research has been conducted to provide evidence for the relationship between self-efficacy and learning strategies, learners' linguistic performance, causal attributions and anxiety. The research conducted on the relationship between self-efficacy and linguistic competence, however, shows a strong focus on reading and listening skills, mostly neglecting the demand for examinations concerning individuals' pronunciation talent. This study incorporates the semantic differential method, an adapted version of the Language Experience and Proficiency Questionnaire, and the Generalised Self-Efficacy Scale to investigate the relationship between 39 university students' self-efficacy and their phonetic aptitude, taking into account their attitude towards near-open front unrounded [æ], their beliefs about their own capability to realise the vowel in a socially acceptable way, and the ratings of 7 native speakers of British English who were provided with 17 random recordings of *The Northwind and the Sun*. The study has shown a range of interconceptual correlations between single attitudes towards near-open front unrounded [æ]. Furthermore, the quantitative analysis has also confirmed that self-efficacy is indeed relatable to individuals' linguistic performance. A significant correlation has also been detected between individuals' phonetic aptitude and their overall pronunciation score as rated by the seven native British English speakers.

D. Leisser (✉)
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria
e-mail: daniel.leisser@univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_5

1 Introduction

"Whether you think you can or whether you think you can't – you're right" (Kroth, 2007, p. 97). This famous quote, which is often ascribed to Henry Ford and is probably one of the most well-known proverbs, does not seem to have much in common with the research conducted in the field of second language acquisition. However, a central discussion currently taking place within this field is concerned with "self-related beliefs" (Mercer, 2011, p. 335) and increasingly with the role of self-efficacy in the successful development of linguistic competence. Along with variables such as "self-concept, self-esteem […] and identity" (Graham, 2007; Henry, 2009; Mercer, 2011, p. 336; Morita, 2004; Norton, 2000; Pellegrino, 2005; Rubio, 2007), self-efficacy may be considered a significant self-related construct. Being of crucial importance in the social cognitive framework (Raoofi, Tan, & Chan, 2012), self-efficacy has been described as the "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Raoofi et al., 2012, p. 60). It is therefore both a valid and necessary question to ask as to which degree the relationship between self-efficacy and linguistic competence is meaningful and how reliably measurements can be conducted.

Plenty of research has been done on the ability of learners to understand the tasks provided (Williams & Burden, 1997), the "learners' beliefs in their own abilities to perform a task" (Bandura, 1997, cited in Raoofi et al., 2012, p. 60), the impact of different learning strategies (Cohen, 1998; O'Malley & Chamot, 1990) and the role of motivation (Csizér & Lukács, 2010; Dörnyei, 2001, 2005; Dörnyei & Ushioda, 2009; Gardner, 2000). Affective variables such as motivation and attitude describe different concepts. Whereas motivation may be a "process by which goal-directed activity is instigated and sustained" (Pintrich & Schunk, 1996, p. 4), the definition of attitude heavily relies on context. Despite the conceptual disagreement between disciplines regarding the notion of attitude, it can hardly be reduced to likes and dislikes (Bem, 1970, p. 14). The concept of language attitude is a central one in the field of sociolinguistics. The idea of phonetic attitude, that is, the extent to which individuals positively or negatively respond to a given sound, has not been discussed in detail. From a linguistic perspective the established distinction between competence and performance (Chomsky, 1965) is of utmost importance in this context as linguists are only able to perceive individuals' performance that is their actual linguistic output. The psychological processes underlying individuals' performance, however, remain largely unknown. Therefore, the investigation of self-efficacy as a possible component of language aptitude is an important step to take in order to introduce social cognitive theory to the field of second language acquisition.

In this paper, a study is introduced which aims to shed light onto the relationship between self-efficacy and phonetic aptitude. In the first section the most significant concepts such as self-efficacy and linguistic aptitude will be clarified to avoid misunderstandings due to the great terminological divergence in the field of attitudinal research. After the clarification and elaboration of these concepts, the second section will then familiarise the reader with the research question which lies at the heart of this project, as well as the various hypotheses by the end of this paper. Subsequently, the third section will describe in detail the empirical data collected, considering the methodology applied in the course of the experiment, the individuals partaking and the results generated. The fourth section finally serves to discuss the study's most significant findings and to outline future research perspectives.

2 Self-Efficacy, Linguistic Competence, Language Aptitude and Language Learning

The linguistic self is constituted by a variety of psychological and linguistic aspects, which have taken many shapes and forms in both psychometrics and applied linguistics. Therefore, an a priori clarification of conceptual entities is necessary since the notion of self-efficacy often leads to terminological misconceptions. Self-efficacy has been defined as "beliefs about [individuals'] capabilities to produce designated levels of performance that exercise influence over events that affect their lives" (Bandura, 1994, p. 71). It may therefore be assumed that self-efficacy is very much relevant to language learning processes.

Language teachers tend to motivate students with the claim that if they only believe hard enough in themselves and their abilities, they will be successful. This bold claim can no longer remain neglected by the scientific community and justifies the call for empirical examination. Bandura (2006) states that "[p]erceived self-efficacy can have diverse effects on motivation, thought, affect, and action" (p. 319). At this point, it is crucial to differentiate between linguistic competence (Chomsky, 1965), which refers to a speaker's underlying knowledge of a given language and the notion of language aptitude. Despite the fact that some use the terms interchangeably, it is important to draw attention to the conceptual differences between them. Language aptitude has most commonly been conceived of as a potential for language learning; it is thus a term describing the holistic acquisition of the language system.

This paper will exclusively focus on the notion of phonetic aptitude, which may be defined as an individual's potential for the acquisition and appropriate production of a language's sound inventory. As phonetic aptitude is assumed to be a variable of predictability, namely to which extent an individual is expected to acquire a given sound inventory, it is indeed purposeful to pose the question as to whether self-related beliefs are in any way related or relatable to this sub-concept of language aptitude. In other words, it is worth investigating whether an individual's self-efficacy with regard to its pronunciation talent has an influence on this individual's phonetic realisation of language. Hence, as evident in the nature of the experiment, both variables chosen are inextricably connected with the element of predictability. This general notion of expected success, which has been discussed earlier, shall now be put to the test with regard to its linguistic dimension. This would not only be valuable in terms of the interdisciplinary research conducted in the fields of psychometrics and applied linguistics, but would also raise new questions as to how, why and when self-related beliefs impact language aptitude.

3 Apropos Aptitude: Research Questions and Hypotheses

In the previous section, focus was placed on the possible relation between self-efficacy and language aptitude. It has been explained why self-efficacy and phonetic aptitude are both relevant and promising variables to investigate. In this section, the study's research questions and respective hypotheses will be presented and discussed. Based on the dependent phonemic variable, that is, the near-open front unrounded [æ], the epistemological expectations towards the experiment will be addressed.

Self-efficacy research is not an entirely new approach to the processes of language learning. Bandura (1997) and Schunk (1991) have already drawn attention to the link between learners' prediction of performance and their actual abilities. Self-efficacy and its development may be understood in relation to the tetralogy of master experience, vicarious experience, social persuasion and physiological states (Bandura, 1997, cited in Raoofi et al., 2012, p. 60). The role of self-efficacy as far as learners' interest, persistence, extent or effort in language learning, the goals set and the use of so-called self-regulated strategies are concerned has been repeatedly examined (Carmichael & Taylor, 2005; Lane, Lane, & Kyprianou, 2004; Linnenbring & Pintrich, 2003; Pajares, 1996, 2003; Raoofi et al., 2012; Schunk, 2003). Some of the research conducted so far has focused on the positivist aspect of grading (Hsieh & Schallert, 2008; Mahyudding et al., 2006; Mills, Pajares, & Herron, 2007) and learners' degrees of proficiency in particular fields of the target language (Raoofi et al., 2012, p. 60). What is more, reading and listening skills have been a central aim in the investigation of language abilities related to self-efficacy (Rahimi & Abedini, 2009; Magogwe & Oliver, 2007; Mills, Pajares, & Herron, 2006, Mills, Pajares, & Herron, 2007; Tilfarlioğlu & Ciftci, 2011). To the best of my knowledge, phonetic aptitude has not been investigated with regard to self-related beliefs such as individuals' self-efficacy. It is indeed true that the measurement of self-efficacy is to a large extent task-specific. Precisely because the pronunciation talent may be related to a variety of cognitive and personality factors (Dörnyei, 2006), the role of both attitude and self-efficacy should no longer be neglected in the scholarly discourse.

4 Method

This study aims to extend our knowledge of the relationship between self-efficacy and phonetic aptitude, focusing on the question as to whether there is a direct relationship between self-efficacy and phonetic aptitude. If so, is there a correlation between participants' psychological attitude towards near-open front unrounded [æ] and their beliefs about their own capability to realise the vowel in a socially accepted way? In the process of hypothesis testing, the following hypotheses are intended to be either proven or rejected:

H_1: There is a direct relationship between individuals' self-efficacy levels and their phonetic aptitude.
H_2: Individuals who show high self-efficacy levels show higher phonetic aptitude results.
H_3: Individuals who show high self-efficacy levels receive better pronunciation ratings by native raters.
H_0: There is no relationship between individuals' self-efficacy levels and their phonetic aptitude.

4.1 Semantic Differential

In the previous section the research questions and the hypotheses which underlie the study were presented. This section will also not only give a detailed account on the methodology employed, but also provide a meticulous description of the quantitative data collected. The experiment is based upon Osgood's semantic differential method (1957) which is also used in other fields of research to measure beliefs and attitudes towards objects. It seems to be a general consensus that beliefs and attitudes do have a strong impact on human behaviour and social interaction (Ferguson & Bargh, 2004; Rosenhan, 1973; Rosenthal, 1985). It appears as though it is particularly challenging to measure the level of self-efficacy and the 'real' connotations individuals associate with phonemic variables such as [æ].

It cannot be denied that participants often show "a tendency to provide socially desirable answers" (Raoofi et al., 2012, p. 66) and researchers must indeed be aware of this fact.

The semantic differential is considered an accurate method to measure individuals' attitudes, especially if it is administered across multiple groups, which show a degree of homogeneity. The semantic differential method may provide valuable insights in the construction and articulation of linguistic realities, namely "how [students] feel and think about learning a new language" (Aragao, 2011, p. 332). According to argumentum a maiori ad minus, it is reasonable that if one may measure students' cognitive and affective attitudes towards a language, the same measuring technique might also apply for individual sounds. The observation that individual sounds are more significant in individuals' social relationships than assumed was confirmed when Tracy, Bainter, and Satariano (2015) were able to empirically account for individuals' tendency to focus on particular phonemes when drawing conclusion about another individual's sexual orientation. It is thus no longer tenable that attitude can only relate to entire language systems and not their underlying sounds.

pleasant	○	○	○	○	○	○	○	unpleasant

Fig. 1 Example of measuring participants' psychological attitude towards [æ]

4.2 Linguistic Socialisation

In order to gain sufficient information on individuals' linguistic socialisation, participants were provided with an introductory questionnaire, which is based on an adapted version of the *Language Experience and Proficiency Questionnaire* (Marian, Blumenfeld, & Kaushanskaya, 2007, pp. 967–968). They then were instructed to complete the semantic differential form, which consists of two sections; both comprising a row of contrasting adjectives in direct juxtaposition with seven circles providing the possibility to express gradable attitude. Section 1 is intended to investigate participants' psychological attitude towards [æ], including numerous adjectives which are intended to measure the participants' emotional response to the sound. An example is given in Fig. 1.

Section 2 focuses on participants' degree of self-efficacy, instructing them to mark the adjectival circles on each scale where they consider most accurate. The closer the circle is to the adjacent adjective, the stronger this adjective applies. Whereas section 1 is heavily based on evaluative adjectives, namely whether [æ] is received as rather desired or undesired, section 2 has been designed to allow the participant to deliver an anonymous categorisation of their beliefs about their own phonetic aptitude when producing [æ]. Despite the fact that the two sections of the semantic differential show a high level of redundancy, face validity has been a major aspect taken into account when the scales were designed. According to Bandura (2006, p. 318), self-efficacy scales "should measure what they purport to measure, that is, perceived capability to produce given attainments". This was especially considered with those items that seem to be prone to semantic fuzziness.

4.3 Generalised Self-Efficacy Scale

As the final stage of self-description, participants were asked to complete the *Generalised Self-Efficacy Scale* (Schwarz & Jerusalem, 1993) to allow for an in-depth analysis of their general level of self-efficacy. Participants' generalised self-efficacy was measured with the following ten items:

1. I can always manage to solve difficult problems if I try hard enough.
2. If someone opposes me, I can find means and ways to get what I want.
3. It is easy for me to stick to my aims and accomplish my goals.
4. I am confident that I could deal efficiently with unexpected events.

5. Thanks to my resourcefulness, I know how to handle unforeseen situations.
6. I can solve most problems if I invest the necessary effort.
7. I can remain calm when facing difficulties because I can rely on my coping abilities.
8. When I am confronted with a problem, I can usually find several solutions.
9. If I am in a bind, I can usually think of something to do.
10. No matter what comes my way, I am usually able to handle it.

The problem of acquiescence has recently been addressed by Danner, Aichholzer, and Rammstedt (2015) and has been taken into account when creating the questionnaires and survey forms. Both the semantic differential method and the Generalised Self-Efficacy Scale may be successful in avoiding participants' "agree[ment] […] to items regardless of content" (Danner et al., 2015, p. 119). The methodology applied therefore focuses on the triangulation of participants' psychological attitude towards the sound in question, their sound-specific self-efficacy levels, and their general self-efficacy. Seventeen participants agreed to being recorded when reading out Aesop's well-known fable *The Northwind and the Sun*. The respective recordings were then played to seven native speakers of British English, who were instructed to grade the 'native-ness' or 'foreignness' of the accents heard and the correct or incorrect realisation of [æ].

5 Participants

5.1 Self-Description: LEAP-Q

The test group consists of 39 individuals who are all students at the University of Vienna. 17.95% of the participants are male, 82.05% are female. 74.4% of the participants study English as part of their teaching degree, 23.00% pursued a bachelor's degree. 2.6% were registered for both university programmes. The mean age of all participants is 23.67. Whereas 10.3% of the participants disclosed a vision problem, there are no individuals with language impairments.

All students partook in the language proficiency course *'Practical Phonetics and Oral Communication Skills 1'* (PPOCS1). In this course, focus is placed on the accurate acquisition of segmental and suprasegmental features of standard British English. In order to fulfil requirements of the final oral exam, participants are obliged to attend the PPOCS1 language lab, which is mainly concerned with the practical reinforcement of the course contents. According to Marzuki (2014, p. 118) a language lab may be conceived of as an "audio or audio-visual installation used as a tool to assist language teaching". For 76.9% of the participants it was the first time that they attended the course PPOCS1, whereas 20.5% had already taken it at an earlier point in time. 2.6% of the participants did not disclose either of the options.

5.2 Self-Ratings: Language Dominance and CEFR

There were no native speakers of English amongst the participants. 84.6% of the participants stated that English was their second language in acquisition. The mean age of onset was 8.78. The mean age of first reading in English was 11.63. The mean age of first fluency in English was 15.56. English was the second most dominant language in acquisition for 82.1% of the participants. 17.9% disclosed English as the third most dominant language in their lives. 18.9% of the participants were not raised in Austria, but settled in the country at an earlier point in time, whereas 82.1% were raised in Austria. According to the Common European Frame of Reference (CEFR) 8.72% of the participants indicated they had B2 proficiency in English. 15.77% rated their own proficiency as C2. 75.51% disclosed that their language proficiency is at C1 level.

5.3 Language Proficiency: Productive and Receptive Language Skills

Participants were instructed to rate their own language proficiency in speaking, understanding spoken English, and reading from 0 to 10. The mean speaking proficiency of the participants is 7.54, their average understanding of spoken English is 8.79 and their alleged reading proficiency lies at 9.00. This leads to the assumption that participants tend to consider themselves less proficient in productive skills, such as speaking, than they do in receptive skills, such as listening or reading. Using the same scale (0–10) for the question as to how much interaction with friends and family, reading, watching TV, listening to the radio and language tapes or self-instruction had contributed to their language learning, participants stated that reading (9.05) and watching TV (7.59) are the most salient factors. These findings are also supported by a high degree of concurrent validity as both reading (8.33) and watching television (7.10) again prove to be the main sources of language exposure.

5.4 Dialectal Preference

When asked to indicate dialectal preferences 87.2% of the participants opted for British English, 5.1% indicated American English. The remaining 7.7% did not indicate either of the optional varieties. The mean strength of dialect preference for British English is 7.03 out of 10. Participants were also instructed to rate how much of a foreign accent they have in English in their own perception. The mean strength of foreign accent is 5.33, which reveals quite some undecidedness amongst participants.

6 Attitudinal Response to Near-Open Front Unrounded [æ]

As previously discussed, the first section of the semantic differential focused on individuals' attitude towards the near-open front unrounded [æ]. Participants were instructed to indicate on a scale from 1 to 7 which of the opposing adjectives best applies. Critics have asserted that it would be impossible for individuals to show positive or negative affective responses to a single sound such as [æ]. In this section, participants' affective response to near-open front unrounded [æ] will be described and discussed.

6.1 Means (Section 1)

It is noteworthy that two participants did not entirely complete the questionnaire, which is the reason for the minimal divergence in the overall number of participants. The adjectival pairs to which participants reacted most strongly are <important> <unimportant> (2.8158), <valuable> <invaluable> (2.8974) and <distant> <close> (2.8974). Pairs such as <pleasant> <unpleasant> (3.9744) did not trigger a strong attitudinal response. There is a significant correlation between the vowel's perceived importance and its perceived utility. This may indeed corroborate the claim that a large number of participants may have expected the vowel to be of good use in summative assessment in the approaching oral exam.

A sociolinguistic analysis (Leisser, 2016) previously conducted on the data showed that social variables such as age and gender do not seem to have an influence on the affective responses triggered by the semantic differential. This is of relevance as critics have claimed that social categories may influence the participants' behaviour when completing the questionnaire.

6.2 Means (Section 2): Vowel-Specific Self-Efficacy

The second section of the semantic differential shows participants' affective response to their own production of the vowel in question. This part seems to account for some concurrent validity with regard to participants' vowel-specific self-efficacy levels, namely how they feel when producing the sound. The comparison of means shows that on average, participants did not feel overly nervous when producing the sound. The data also shows that participants tend to feel strong-willed (5.0513), effective (3.2105), skilful (3.2105), energetic (3.3158) and generally able (3.3421). Cronbach's alpha (0.827) shows satisfactory results with regard to the semantic differential's overall internal consistency.

There is high degree of fluctuations, which reveals an overall tendency to high sound-specific self-efficacy levels. However, it is a central question which concepts may be formed based on the adjectival descriptions of the attitudinal or affective response participants disclosed. It may be argued that even slight fluctuations may be considered meaningful since due to the setting participants are very likely to be tentative with their responses. On average participants show a tendentiously positive response to the sound, which may also relate to the assessment at the end of term.

The descriptive statistical results have shown that participants' attitude towards the phonetic feature and their self-perceived self-efficacy when producing the sound shows a considerable degree of variation. Individuals' generalised self-efficacy ranged from a minimum of 23.00 to a maximum of 38.00. The highest possible level of self-efficacy according to the generalised self-efficacy questionnaire is 44.00. The mean level of self-efficacy of all participants was 30.71 according to the generalised self-efficacy scale. It will later be discussed whether participants' attitude towards the sound, their sound-specific self-efficacy and their generalised self-efficacy are by any means related. There is no age or gender-related divergence in individuals' generalised self-efficacy which would be a possible objection to the data presented, as gender differences in particular are often assumed.

7 Accent Ratings and Scores

7.1 Subgroup

Sixteen female and one male participant (N = 17) agreed to be recorded after a language laboratory session whilst they were reading Aesop's fable *The Northwind and the Sun*. Accordingly, 94.1% of this subgroup of participants were female, and 5.9% were male. The mean age of the group was 24.53. The participants stated that on average they had been learning English for 11.71 years. The data also shows that the mean age of language onset was 9.26. On average, participants indicated that they started reading English at the age of 12.11 and were able to read English texts fluently at the age of 15.50. Furthermore, they enclosed that they were fluent in English at the age of 16.06. On a scale from 0 to 10, participants rated their English reading competence on average 9.06, their ability to understand spoken English 8.82 and their speaking skills 7.41. 58.8% rated their language proficiency as C1 according to the Common European Frame of Reference, whereas 9.1% stated they had only reached B2, and 12.1% claimed C2 proficiency. 23.5% ticked they had already attended the university course Practical Phonetics and Oral Communication Skills 1, 76.5% said they had not. None of the participants indicated that they were native speakers of English.

7.2 The Northwind and the Sun

As previously mentioned, the text that participants were instructed to read was taken from Aesop's fables and is named *The Northwind and the Sun*. For this study, the following version was used:

> The North Wind and the Sun were disputing which was the stronger, when a **traveller** came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the **traveller** take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the **traveller** fold his cloak around him; and at last the North Wind gave up the attempt. Then the Sun shined out warmly, and immediately the **traveller** took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two. (International Phonetic Association, 2003, p. 44, my emphasis)

The recordings were then rated by seven native speakers of British English, three being female, four male. The mean age of raters was 33.71. All of them confirmed that they were native speakers of British English. The raters did not only consider the apparent strength of the accent they heard, but also how often they could hear [æ] in the word 'traveller' produced correctly. It may be noted that the lexical item 'wrapped' also includes the vowel [æ]. However, it was not included in the ratings in order to increase the level of inter-rater reliability and make the findings more comparable. When designing the rating scales the leading motive was not only to make findings comparable, but also to ensure that face validity is very high. To achieve this aim, one cannot find specific explanations or jargon in the instruction sheet. When pre-testing the rating scales, I encountered the potential issue that asking participants to indicate a specific sound such as [æ] might lead to considerable confusion as to what exactly is expected of them. For this reason, the instruction sheet was adapted and asks for the letter 'a' in the word traveller. Whilst it is clear that this is not an appropriate representation of a sound, especially among theoretical phoneticians, this sacrifice was essential in order to ensure that raters are fully aware of the task at hand. Finally, the following wording has been chosen:

> Please use the accent rating scale on the following pages to rate the native or non-native voices you are about to hear. The rating scale consists of two poles: 'no foreign accent' and 'very strong foreign accent'. Please put a cross in the circle which you consider most appropriate for each speaker. The closer the circle is to either of the poles, the stronger it applies. Please also mark how often the letter 'a' in the word 'traveller' is produced appropriately.

Whereas raters were not always confident about the overall accent rating, they could easily distinguish between what they perceived as correct or incorrect realisations of near-open front unrounded [æ]. The mean of all average ratings is 2.75, which indicates that participants' success rate lies between two and three acceptable articulatory productions. Pearson's correlation shows a negative dependency between the incorrect production of [æ] and low accent score levels. This increases the level of validity as raters' judgement with regard to participants' accents is supported by their perception of the correct and incorrect productions of [æ]. In the following section, the various correlations between participants' phonetic attitude, their self-perceived self-efficacy, their generalised self-efficacy and the accent ratings provided by the seven native speakers will be examined.

8 Results: Concepts and Correlations

Before the correlations of the three main variables are introduced, this section will provide the most important correlations between the adjectival pairs in the semantic differentials. These do not only confirm that there is remarkable convergence between the positive and negative concepts which may be derived from the scales, but also that the findings are indeed meaningful.

8.1 Semantic Differential: Attitudinal Correlations

To give but two examples, there is a significant correlation between the vowel's perceived valuableness and its perceived pleasantness. Furthermore, there is a highly significant correlation between the helpfulness perceived and its attainability. There are also correlations between participants' attitudinal response to the sound and their sound-specific self-efficacy, e.g. between their perceived capability and the sounds' pleasantness or the degree of self-confidence and the extent to which they are familiar with the vowel.

8.2 Semantic Differential: Self-Efficacy and Accent Scores

There is doubtlessly a meaningful relation between individuals' attitude towards the phonetic feature and their perceived self-efficacy when producing it. Participants' generalised self-efficacy does not correlate with the attitudinal response of the semantic differential. There is however a significant, negative correlation between individuals' generalised self-efficacy and the strength of accent perceived by the raters. This seems to suggest a relationship between individuals' generalised self-efficacy and their linguistic performance, but could also account for a relationship between self-efficacy and phonetic aptitude in general. It seems that the higher individuals' generalised self-efficacy levels were, the better they scored on their native accent ratings. The following Figs. 2 and 3 show the correlations between individuals' perceived generalised self-efficacy, strength of accent and native ratings for [æ] with the respective t-test results.

Correlations

		efficacy	strength_accent	ae_corr
efficacy	Pearson Correlation	1	-,579*	,336
	Sig. (2-tailed)		,015	,187
	N	17	17	17
strength_accent	Pearson Correlation	-,579*	1	-,575*
	Sig. (2-tailed)	,015		,016
	N	17	17	17
ae_corr	Pearson Correlation	,336	-,575*	1
	Sig. (2-tailed)	,187	,016	
	N	17	17	17

*. Correlation is significant at the 0.05 level (2-tailed).

Fig. 2 Correlations between perceived generalised self-efficacy, strength of accent and native ratings for [æ]

One-Sample Test

	Test Value=0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
efficacy	30,604	38	,000	29,92308	27,9437	31,9024
strength_accent_mean	12,411	16	,000	5,01681	4,1599	5,8737
ae_corr_mean	21,632	16	,000	2,74790	2,4786	3,0172

Fig. 3 T-test between perceived generalised self-efficacy, strength of accent and native ratings for [æ]

8.3 Self-Rating and Native Rating

After the assumed relationship between individuals' generalised self-efficacy and their overall accent score was found to be confirmed, the relationship between individuals' self-ratings and the ratings provided by the native speakers was investigated. There seems to be a highly significant correlation between these two variables which again contributes to the data's validity, which is also supported by the high internal consistency within native ratings of accent scores (Cronbach's alpha = 0.847).

The correlation reveals that individuals' personal accent ratings and their native raters' judgement differ only slightly, which suggests that individuals are able to rate their own accents reliably. However, so far no direct relationship could be found between individuals' generalised or vowel-specific self-efficacy and their accent scores. Therefore, it is necessary to examine individuals' linguistic socialisation to find possible explanations for the given correlations, and for the fact that there is no sufficient correlation between individuals' generalised self-efficacy and their vowel-specific self-efficacy. Other variables such as the age of language onset, the age of first fluency in speaking and reading do not correlate with individuals' generalised or vowel-specific self-efficacy.

9 Discussion

This section will discuss the findings presented and attempt to provide explanations why H_1 seems to be correct, namely that there is a relation between individuals' self-efficacy levels and their phonetic aptitude. Students with high self-efficacy levels seemingly show higher phonetic aptitude than their peers. It may however also reflect the conceptual issues which arise in interdisciplinary work between the domain of social psychology and that of linguistics. The study presented was also intended to provide new insights in the measurement of self-concepts, such as self-efficacy and the linguistic variable of phonetic aptitude, which has been defined as a potential rather than a fixed state. It is hard to deny that "affective attitudes, just like cognitive attitudes, influence speech perception" (Nguyen, Shawa, Tyler, Pinkus, & Best, 2015, p. 4). Following Podesva's (2006) view that phonetic detail has much been neglected, this study intended to fill this research gap and investigate whether individuals are able to develop affective attitudes towards an individual sound and whether these attitudes are by any means related to their phonetic aptitude.

9.1 Negotiating the Meaning of Attitude and Vowel-Specific Self-Efficacy

Section 1 of the semantic differential chart shows that individuals are indeed able to form attitudes towards individual sounds. This was also expected as the setting of the experiment is the language laboratory complementing a university course in phonetics and phonology. Near-open front unrounded [æ] was tendentiously perceived as valuable (2.8974), exclusive (3.3077), useful (3.3333), attainable (3.5641), annoying (3.4872) and important (2.8158). As previously mentioned, the correlation found between the perceived importance and the sound's utility may be interpreted in the context of summative assessment at the end of the phonetics course. Critics of the semantic differential method have repeatedly voiced their concern regarding the construct validity of the data generated. The use of semantic differentials in attitudinal research is indeed accompanied by methodological discussions on the interpretation of the findings. Al-Hindawe (1996, p. 8) has drawn attention to the relationship between the types of elicitation and the degree to which the data is "analysable, meaningful and relevant to the study at hand". The semantic differential method therefore challenges researchers' abilities to make sense of the numerical values when scoring the answer sheets.

In the case of attitudes towards sounds this is indeed a complex matter. Nevertheless, it is certainly possible to derive meaningful conclusions from the scales used in this study. A comparative analysis of both sections reveals that there are a number of highly significant correlations between individuals' attitude towards [æ] and their vowel-specific self-efficacy, which indicates that there may be a connection between the attitude individuals hold towards a linguistic item and the item-specific self-efficacy. The study has shown that the items which were intended to capture the degree of item-specific self-efficacy, such as <weak-willed> <strong-willed> (5.50513) <effective> <ineffective> (3.2105) <skilful> <unskilful>> (3.2105) <energetic> <lethargic> (3.3158) and <able> <unable> (3.3421) show meaningful results. Individuals' vowel-specific self-efficacy could therefore be determined, which is also a result of the high redundancy across the items used. Bandura (2006, p. 308) argues that "[e]fficacy items should accurately reflect the construct". The notion of content validity was a crucial aim in the construction of the semantic differential scales. This was also inextricably linked to the scales' face validity as the degree to which individuals are able to complete the scales satisfactorily heavily depends on how their "perceived capability" (Bandura, 2006, p. 308) is elicited.

The sociolinguistic analysis conducted before the introduction of the concept of self-efficacy has shown that there is no sufficient evidence to claim that there is a gender-related divergence between individuals in terms of their attitudinal response. The socio-phonetic dimensions of the data generated also raises important questions with regard to the current developments in the field of socio-phonetics, namely how the attitudes individuals hold towards certain sounds and social meaning-making processes are related. Second language learning is certainly embedded in the sociolinguistic context and must therefore not be excluded from such contemplation.

There has been no evidence to suggest that there is a relationship between individuals' generalised self-efficacy, as measured with the generalised self-efficacy scale, and their vowel-specific self-efficacy. Individuals who scored high on self-efficacy on the generalised self-efficacy scale did not necessarily do so in section 2 of the semantic differential.

9.2 Accent Rating Scores

The study shows a negative correlation between individuals' generalised self-efficacy and the strength of accent perceived by the raters. In other words, individuals with high generalised self-efficacy levels were more likely to receive lower accent scores. As previously mentioned this could account for a valid relationship between self-efficacy and phonetic aptitude. It certainly confirms that self-efficacy is relatable to individuals' linguistic performance. Prescriptively speaking, individuals with high generalised self-efficacy levels have performed 'better' on the task at hand than those with low generalised self-efficacy levels. There is also a relationship between individuals' self-ratings in terms of accent strength and their native raters' judgement. Individuals who considered their accent rather strong received higher accent scores when rated by native speakers of British English. This seems to prove that individuals' self-ratings may yield valid statistical results. Not only does individuals' generalised self-efficacy seem to affect their linguistic performance, individuals appear to be able to rate their own accent relatively accurately. This seems to confirm Bandura's (2006, p. 319) claim that perceived self-efficacy can indeed influence affect and action, which may also hold true for individuals' linguistic performance. The notion of accent accordingly shifts from a stable feature of spoken discourse to a variable which can be modulated. Self-related beliefs such as perceived self-efficacy may indeed influence this modulation of accent.

9.3 Attitude, Aptitude and Self-Efficacy

As this study has shown, there are interconceptual correlations between individual attitudes towards linguistic features such as near-open front unrounded [æ]. These phonetic attitudes may have an influence on learners' self-concepts and affect learners' success. Denissen, Eccles, and Zarett (2007, p. 430) have drawn attention to a number of studies investigating the relationship "between liking a subject, and doing well in the subject" (e.g. Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002; Renninger, 2000). In order to conduct systematic and reliable attitudinal research, the variables with which analyses are undertaken need to be meticulously defined. After all, the semantic differential only provides data within a specific psychological domain. This is particularly relevant if the subject of investigation is the relationship between linguistic attitudes and the linguistic self with its various similar, psychological manifestations. The central question as far as attitude is concerned must therefore relate to the impact that attitude has on individuals' self-concepts. This study shows that attitude and self-efficacy may be related. Individuals who regarded near-open front unrounded [æ] as familiar would also describe themselves as self-confident when producing the sound. This observation is relevant as Brophy (1998) relates self-efficacy to individuals' tendency of showing a higher degree of persistence when facing difficulties in the language learning process. Another significant aspect may be the role of attitude in the psychological construction of self-concepts such as self-efficacy. Matthews (2010) states that "[s]tudents with low self-efficacy for a given topic may devalue the domain, may elect not to participate or invest effort in learning, and may consider themselves lacking in the ability to succeed in it" (p. 619). Despite the fact that the number of correlations between attitudinal variables and those of vowel-specific or generalised self-efficacy is indeed rather small, it is necessary to discuss the relationship between attitude and self-efficacy on the level of theoretical propositions.

Agreeing with Barcelos (2003), Woods (2003), and Mercer (2011) argues that one of the greatest difficulties in the investigation of individuals' beliefs is that they "are notoriously difficult to define" (p. 336). It seems that self-efficacy beliefs do not constitute an exception to this statement. The relation to individuals' attitudes towards the sound under examination must therefore always be scrutinized with respect to the context and the setting of the university course. Woodrow (2011) maintains that "[s]elf-efficacy differs from other conceptualisations because it is domain specific and focuses on a specific action" (p. 512), e.g. the production of a correct sound in summative assessment. This is also a problem in terms of the validity of findings as individuals' beliefs are not merely "static mental representations" which are not affected by time or context (Mercer, 2011, p. 336). Hence, the relationship between linguistic attitudes and self-concepts such as self-efficacy must be investigated from multiple angles to ensure the validity and reliability of the data collected.

Much like the beliefs and attitudes examined, the notion of phonetic aptitude describes a potential for linguistic ability rather than a fixed state. Sternberg (2002) similarly describes language aptitude as a concept "that involves multiple aspects" and cannot be reduced to "some single fixed quantity" (p. 14). This is a central assumption underlying the research conducted as the data only leads to inferences on what individuals produced at a given point in time. The notion of phonetic aptitude, however, goes beyond the Chomskyan construct of linguistic competence. The investigation of phonetic attitude must necessarily draw on individuals' performance which is the basis for any further contemplation. The spoken text produced spontaneously raises a number of questions with regard to self-efficacy beliefs and attitudes. If individuals were able to adapt their psychological self-concepts, would this not have an impact on their general linguistic aptitude, possibly including their phonetic encoding or decoding abilities? This is particularly interesting in the context of the "complex human motor skills" (Levelt, 1989, cited in Hu et al., 2013, p. 366), which seem to limit individuals' possibilities in second language learning.

Language aptitude, self-concepts and attitudes are dynamic and prone to change. The study presented has shown that what some in the scientific community conceive of as language aptitude may be even more complex on a conceptual level. Dörnyei (2006) has linked individuals' pronunciation talent to a variety of factors, including the domains of cognition and personality. One may hence raise the question to which extent self-concepts such as self-efficacy, and cognitive or affective attitudes contribute to the configuration of language aptitude. In the context of a language laboratory this is particularly relevant as individuals are expected to acquire an entire sound inventory of a given language within a very limited timeframe. Accordingly, individuals with very high phonetic aptitude would excel and those with very low phonetic aptitude would yield rather low results. Nevertheless, the pressure which individuals are facing seems to positively affect their academic success since the majority seems to meet the challenge successfully. The findings of this study may also be relevant to the investigation of individuals' mental concepts with respect to the "level of skill in the attainment of education goals" (Denissen et al., 2007, p. 430).

If language aptitude is co-determined by attitudes and self-beliefs, it seems necessary to conduct similar studies across various education systems. It is significant how language aptitude is constituted in individuals and which external factors may have an impact on it. Phonetic imitation ability appears to be distributed unevenly across the human species as "aptitude, ability, and success in sound imitation learning" differ greatly from individual to individual (Reiterer, Hu, Sumathi, & Singh, 2013, p. 1). Furthermore, scholarly discourse has seen a crucial distinction with regard to the subfields of language aptitude. Research has shown that individuals display a tendency towards either grammatical or phonetic aptitude (Reiterer et al., 2011). It may therefore be necessary to investigate and compare the various settings of language didactics, e.g. in language laboratories, which are intended to increase the effect of students' practice whilst fostering language teachers' productivity (Marzuki, 2014). At any rate, language aptitude seems to be undergoing a conceptual metamorphosis which must not be neglected in studies currently conducted.

The constitution of the linguistic self and its various psychological constituents should be an essential component of aptitude research. The construction of appropriate and meaningful questionnaires may be particularly challenging since self-concepts such as self-efficacy demand a very precise approach to the psychometric measurement required. Nevertheless, the investigation of the relationship between self-efficacy and productive language abilities, such as phonetic coding abilities, is of utmost importance to provide insights into the constitution of the linguistic self and its psychological sub-concepts. If language aptitude was biologically determined, the observation that "highly efficacious students are confident about what they can achieve, [that] they set themselves challenges and are committed to accomplish them, [and that they] work harder to avoid failure" (Ching, 2002, cited in Yilmaz, 2010, p. 683) would hardly be meaningful. Therefore, both attitudinal as well as self-concept research is necessary to clarify to which extent individuals are able to control what seems to fuel their linguistic abilities.

Another aspect which could not be elaborated on in this project is the role of anxiety in language learning processes such as a phonetics course. The role of self-efficacy as a component of language aptitude may be contextualised with other psychological factors, one of which being the notion of "lathophobia" or "error neurosis" (Bolitho, 2011a). It is indeed reasonable to argue that an individual's self-efficacy may also be influenced by other factors which are crucial in the process of developing linguistic competence. Following Bolitho's (2011a) approach, this paper argues for an in-depth investigation of the relationship between individuals' attitudes towards a specific domain, e.g. phonetic encoding ability, their domain-specific self-efficacy and the psychopathological notion of pronunciation angst. Courses in phonetics which require the attendance of a language lab may contribute to an increase of students' self-efficacy levels. They may, however, also have a negative impact on the formation of students' linguistic self as "[o]f all the language systems, phonology is the one most closely associated with identity, with who we are and how we feel about it" (Bolitho, 2011b). It therefore seems reasonable to call for a turn to affect, to incorporate and link self-concepts and pronunciation angst. Self-efficacy is likely to have a direct impact on individuals' linguistic performance in pronunciation. The central question with respect to its other conceptual siblings such as anxiety is whether self-efficacy can ever be measured as a distinct psychological category or whether this concept remains a highly relative variable.

In other words, is it possible to measure self-efficacy without measuring other relevant concepts? In the field of language aptitude research, this question raises another one, namely which place attitudinal or emotional categories have taken in language aptitude so far.

In his discussion of the *Modern Language Aptitude Test*, Singleton (2014, p. 560) points out that the common language aptitude tests have been criticised by linguists since such tests do not take psychological concepts or emotional aspects into consideration, referring to Stansfield (1989, pp. 3–4) and Parry and Stansfield's (1990) criticism:

> The aptitude tests currently in use […] do not take into account new insights […] on the human learning process in general and on the language learning process in particular. Nor do they take into account […] the relation of attitudes, motivation, personality, and other emotional characteristics and predispositions to second language learning. (Singleton, 2014, p. 560)

This study has shown how language aptitude may be related to such attitudes and emotional characteristics. It would be far-fetched to claim a direct relationship between language aptitude and individuals' self-efficacy levels. Nevertheless, the assumption is hardly undeniable that the measurement of language aptitude without the consideration of any cognitive or affective attitudes, or the definition of self-concepts such as self-efficacy, can only lead to an abridgement of scientific recognition. As previously mentioned, the main question here is to which extent attitudinal or emotional categories influence individuals' linguistic development, be it in the field of second language acquisition or in an institutionalised context such as a language laboratory. Chan and Lam (2010) have discussed the relationship between students' self-efficacy and the feedback provided by their teachers. Analogically, this observation may also be discussed with regard to other social interactions involving parents or peers.

Based on the detailed LEAP-Q, this study enables researchers to examine individuals' linguistic socialisation. There may not be sufficient evidence at this point to argue for a correlation between individuals' linguistic socialisation and their self-efficacy level. However, the linguistic self is a complex, psychological configuration which may also be understood with the help of a psychic determinism. According to a broader understanding of this psychoanalytical concept, language learning, as an essential aspect of human cognition, would be co-determined by psychological processes. The narrower approach to this psychic determinism assumes that human beings are not necessarily fully in control of their own behaviour (Kramer, 2009, p. 237), which might also apply to the context of second language acquisition. This raises questions with regard to the relationship between individuals' linguistic socialisation and their potential for language learning. To which extent are assumptions on the role of individuals' linguistic development meaningful for predictions about their language aptitude at a later point in life? Is it possible to measure self-concepts and relate the findings to individuals' linguistic socialisation? In this study, the age of second language onset did not correlate with any of the attitudinal variables, nor was there sufficient evidence to argue for a link between individuals' exposure to their L2 and self-concepts such as self-efficacy. It is nonetheless crucial to investigate the relationship between individuals' linguistic experiences and their self-concepts.

This may particularly be relevant to the discussion on the origins and functions of mental mechanism (Nesse & Lloyd, 1992) which lie at the basis of language learning mechanisms. Evolutionary approaches in the domain of psychology may be helpful to redefine the notion of language aptitude with regard to the fitness of the language learning strategies. Nowak, Plotkin, and Krakauer (1999, p. 147) ask the provocative question as to why human beings learn language signals at all, and why they are not simply inherited. In the context of self-concepts and language aptitude, this is indeed an important thought as there is still considerable disagreement on the roots of human language and on the question as to which evolutionary interpretation should be ascribed to cognition and affect in language learning processes. Thus, future interdisciplinary research ought to focus on the directedness and purposefulness of human language. Next to the semantic encoding process, phonetic aptitude may be considered one of the most essential aspects of the evolution of language. Individuals who are convinced of their own abilities and their chance to attain a pre-defined goal, in other words, those who have high self-efficacy are simply fitter than their peers and are therefore very likely to pass on their genes. Tecumseh, Hauser, and Chomsky (2005) have pointed out that

> questions about original function are of a different logical type. It is an unfortunate fact that the two main sources of data to address such historical issues, namely paleontological and comparative, are simply unavailable for behavioral traits unique to one species. (p. 185)

Even though the so-called original function of human language itself is debatable, it is a fact that it is indeed highly functional and that individuals' language aptitude is not only a variable of potential, but also of evolutionary fitness. Future research must hence aim to expand our knowledge of the interplay between language learning processes and the psychological notions of cognition and affect. This study has shown a positive correlation between individuals' generalised self-efficacy and their total accent score. Individuals with higher self-efficacy were more likely to be considered less 'foreign' than their peers with lower self-efficacy. This could be interpreted in the context of evolutionary phonetics and phonology, which may offer interesting perspectives on the subject matter. Speaking in terms of evolution, a high phonetic coding ability would certainly prove advantageous in the process of natural selection as it enables individuals to become and remain part of social groups. Although it seems difficult to identify an exact stage in human development when the seed of self-concept started growing, it appears to be a fruitful undertaking to conduct longitudinal studies on the linguistic socialisation of individuals, starting before second language onset. In doing so, data on psychological variables and concepts can be collected at the same time whilst individuals' language aptitude is tested. This may possibly shed light onto the "selective processes [that] drove the evolution of the speech system" (Knight, Studdert-Kennedy, & Hurford, 2000, p. 8).

Evolutionary thinking is of course not the only approach employed to describe the relationship between the linguistic self and language aptitude. However, it does provide a firm basis for further contemplation for the purpose of human language. It may also prove a useful theoretical framework to examine the functional role of attitude and self-concept in language learning. Hamacheck's (1992) discussion of social contextualisation of the self also seems very relevant to the relationship between attitude, which is directed at external objects, and aptitude as an internal potential for language learning. The data collected in this study seem to show that individuals are able to develop strong attitudes towards individual phonetic items. Even if this may be influenced by the setting in which the study was conducted, claims that individuals are not able to respond meaningfully to an individual sound are no longer tenable. New language aptitude tests are required to allow for an even higher degree of triangulation between phonetic attitude, aptitude and the diverse manifestations of human self-concept. The consideration of evolutionary principles may lead to important discussions on the psychological and linguistic configuration of the self. The role of self-efficacy in the acquisition of phonetic items is but one instance of the unscrambling of language aptitude. It is however an important step towards resolving "one of science's great remaining mysteries" (Knight et al., 2000, p. 1).

10 Conclusion

In this paper a study was introduced which aimed to shed light onto the relationship between individuals' perceived self-efficacy and their phonetic aptitude. The study built on the theoretical framework of social cognitive theory, in which it is assumed that individuals are not merely passive recipients of external factors, but are able to actively have an impact on their environment. A growing body of research has specifically been concerned with the relationship between self-efficacy and learning strategies, linguistic performance, causal attribution and anxiety. The studies investigating the relationship between individuals' self-efficacy and linguistic competence, however, predominantly aim at reading and listening skills whilst at the same time neglecting the need for examination of individuals' pronunciation talent. Therefore, this study was intended to fill the void and contribute recent data to the scholarly discourse.

In this paper focus was not placed on the Chomskyan notion of linguistic competence, but rather on the concept of phonetic aptitude, with particular emphasis on individuals' pronunciation talent. As discussed in this paper, phonetic aptitude is a variable of predictability, that is, how likely an individual is to acquire a given sound inventory. In this context it has been discussed whether self-related beliefs such as self-efficacy are in any way related or relatable to phonetic aptitude. The motivation for this study was hence to extend our knowledge on the relationship between self-efficacy and phonetic aptitude, also taking into consideration individuals' attitude towards the phonetic item under examination.

Following a tripartite structure of data collection, 39 participants completed the Language Experience and Proficiency Questionnaire, two semantic differential forms and the generalised self-efficacy scale. Subsequently, 17 participants who had agreed to be recorded whilst reading Aesop's fable *The Northwind and the Sun* were rated by seven native speakers of British English. The research question underlying the investigation has been whether there is a direct relationship between individuals' self-efficacy and their phonetic aptitude. Four hypotheses were intended to provide possible answers to this question:

H_1: There is a direct relationship between individuals' self-efficacy levels and their phonetic aptitude.
H_2: Individuals who show high self-efficacy levels show higher phonetic aptitude results.
H_3: Individuals who show high self-efficacy levels receive better pronunciation ratings by native raters.
H_0: There is no relationship between individuals' self-efficacy levels and their phonetic aptitude.

The data collected shows that there is sufficient evidence to argue for a direct relationship between phonetic aptitude and self-efficacy. Participants with higher self-efficacy were perceived as less 'foreign' by native speakers of British English. The measurement of vowel-specific self-efficacy, as conducted in section 2 of the semantic differential scales, did not generate sufficient evidence to support the hypothesis that individuals with higher vowel-specific self-efficacy are able to produce near-open front unrounded [æ] in a way more acceptable to L1 speakers. In section 1, a number of correlations were found between single attitudinal items such as the vowel's perceived valuableness and its perceived pleasantness. Whereas the latter did not elicit a strong attitudinal response, items to which participants reacted most strongly were found to be perceived importance (2.8158), perceived valuableness (2.8974), and perceived distance (2.8974). There is a highly significant correlation between the vowel's perceived utility and its importance, which may be interpreted in the context of inevitable summative assessment.

The quantitative analysis of section 2 showed that participants showed a tendency to feel strong-willed (5.0513), effective (3.2105), skilful (3.2105), energetic (3.3158) and generally able (3.3231), which are all items describing vowel-specific self-efficacy items. The comparison of means has revealed that participants were more likely to disclose that they were 'self-confident' when producing the vowel if they also consider the sound 'familiar'.

The generalised self-efficacy questionnaire displays a mean self-efficacy of 30.71 out of 44. Individuals' generalised self-efficacy was normally distributed, ranging from a minimum of 23.00 to a maximum of 38.00. No divergence in the social variables of age and gender could be found. One positive correlation could be detected between participants' perceived self-efficacy and their accent scores. Individuals with higher self-efficacy levels received significantly lower accent ratings than their peers with lower self-efficacy. No correlation could be found between individuals' self-efficacy and their realisation of near-open front unrounded [æ]. Individuals' perceived accent in English strongly correlated with the ratings provided by L1 speakers of British English, contributing additional validity to the data collected. The acceptable or unacceptable realisation of the vowel also correlated with the total accent score.

To conclude, the data shows sufficient evidence for a direct relationship between individuals' self-efficacy and their phonetic aptitude. The investigation has proven that individuals are indeed able to provide reliable self-ratings of their own accent in a foreign language and that they are not only able develop attitudes towards that variety, but also to specific phonetic items, which critics had considered impossible. The data shows that there is a relationship between individuals' self-efficacy and their linguistic performance. Thus, future research must integrate variables such as self-concept and attitude to broaden the contemporary notion of language aptitude as a potential of language learning. The evolutionary interpretation of phonetic aptitude raises the question as to how psychological concepts and language acquisition are linked and how this connection can be purposeful in the selection process. It appears as though the concept of language aptitude may be in need of reconsideration, especially with regard to the various psychological manifestations of the linguistic self. It is therefore indeed reasonable to call for an exhaustive interdisciplinary study to investigate the seemingly peripheral conceptual areas of language aptitude to extend our understanding of the psychology of language acquisition and decipher the psychological components of language aptitude.

References

Al-Hindawe, J. (1996). Considerations when constructing a semantic differential scale. *La Trobe Working Papers in Linguistics, 9*, 41–58.

Aragão, R. (2011). Beliefs and emotions in foreign language learning. *System, 39*(3), 302–313. https://doi.org/10.1016/j.system.2011.07.003

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior 4* (pp. 71–81). New York: Academic.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman and Company.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In T. Urdan & F. Pajares (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307–337). Greenwich, CT: Information Age Publishing.

Barcelos, A. M. F. (2003). Researching beliefs about SLA: A critical review. In P. Kalaja & A. M. F. Barcelos (Eds.), *Beliefs about SLA: New research approaches* (pp. 7–33). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bem, D. J. (1970). *Beliefs, attitudes, and human affairs*. Belmont, CA: Brooks/Cole.

Bolitho, R. (2011a). Teacher-induced neuroses. *British Council*. https://www.teachingenglish.org.uk/article/teacherinduced-neuroses. Accessed 10 Feb 2016.

Bolitho, R. (2011b). Teacher-induced neuroses: Pronunciation angst. https://www.teachingenglish.org.uk/blogs/rod-bolitho/teacher-induced-neuroses-pronunciation-angst. Accessed 10 Feb 2016.

Brophy, J. (1998). Failure syndrome students. In *Clearinghouse on elementary and early childhood education*. Eric Digest. Champaign, IL: University of Illinois, ERIC.

Carmichael, C., & Taylor, J. A. (2005). Seeing self as others see you: Variability in self-efficacy ratings in student teaching. *Teachers and Teaching: Theory and Practice, 15*(5), 541–561.

Chan, J. C. Y., & Lam, S. (2010). Effects of different evaluative feedback on students' self-efficacy in learning. *Instructional Science, 38*(1), 37–58.

Ching, L. C. (2002). Strategy and self-regulation instruction as contributors to improving students' cognitive model in an ESL programme. *English for Specific Purposes, 13*, 261–289. https://doi.org/10.1016/S0889-4906(01)00008-4

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Cohen, A. D. (1998). *Strategies for learning and using a second language*. Essex, UK: Longman.

Csizér, K., & Lukács, G. (2010). The comparative analysis of motivation, attitudes and selves: The case of English and German in Hungary. *System, 38*(1), 1–13. https://doi.org/10.1016/j.system.2009.12.001

Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality, 57*, 119–130. https://doi.org/10.1016/j.jrp.2015.05.004

Denissen, J. J., Eccles, J., & Zarett, N. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development, 78*(2), 430–447. https://doi.org/10.1111/j.1467-8624.2007.01007.x

Dörnyei, Z. (2001). *Motivation strategies in the language classroom*. Cambridge, UK: Cambridge University Press.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*(1), 42–68. https://doi.org/10.1075/aila.19.05dor

Dörnyei, Z., & Ushioda, E. (2009). *Motivation, language identity and the L2 self*. Clevedon, UK: Multilingual Matters.

Ferguson, M. J., & Bargh, J. A. (2004). How social perception can automatically influence behaviour. *Trends in Cognitive Sciences, 8*(1), 33–39.

Fitch, W. T., Hauser, M. D., & Chomsky, N. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, 97(2), 179–210.

Gardner, R. C. (2000). Correlation, causation, motivation and second language acquisition. *Canadian Psychology, 41*(1), 1–24. https://doi.org/10.1037/h0086854

Graham, S. (2007). Learner strategies and self-efficacy: Making the connection. *Language Learning Journal, 35*(1), 81–93. https://doi.org/10.1080/09571730701315832

Hamacheck, D. (1992). *Encounters with the self* (4th ed.). Fort Worth, TX: Harcourt Brace Jovanovich College.

Henry, A. (2009). Gender differences in compulsory school pupils' L2 self-concepts: A longitudinal study. *System, 37*(2), 177–193. https://doi.org/10.1016/j.system.2008.11.003

Hsieh, P. H. P., & Schallert, D. L. (2008). Implications from self-efficacy and attribution theories for an understanding of undergraduates motivation in a foreign language course. *Contemporary Educational Psychology, 33*(4), 513–532. https://doi.org/10.1016/j.cedpsych.2008.01.003

Hu, X., Ackermann, H., Martin, J., Erb, M., Winkler, S., & Reiterer, S. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: Behavioural predictors and neural substrates. *Brain and Language, 127*(3), 366–376. https://doi.org/10.1016/j.bandl.2012.11.006

International Phonetic Association. (2003). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge, UK: Cambridge University Press.

Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development, 73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Knight, C., Studdert-Kennedy, M., & Hurford, J. R. (2000). *The evolutionary emergence of language: Social function and the origins of linguistic form*. Cambridge, UK: Cambridge University Press.

Kramer, G. (2009). *Clinical psychology*. Somerset, NJ: Pearson.

Kroth, M. S. (2007). *The manager as motivator*. Westport, CT: Praeger Publishers.

Lane, J., Lane, A., & Kyprianou, A. (2004). Self-efficacy, self-esteem, and their impact on academic performance. *Social Behaviour and Personality, 32*(3), 247–256. https://doi.org/10.2224/sbp.2004.32.3.247

Leisser, D. H.. (2016). *On phonetic attitude and the linguistic self: Revisiting socio-phonetic perspectives on age and gender*. Unpublished paper. University of Vienna.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT.

Linnenbrink, E., & Pintrich, P. (2003). The role of self-efficacy beliefs in student engagement and learning in the classroom. *Reading and Writing Quarterly, 19*(2), 119–137. https://doi.org/10.1080/10573560308223

Magogwe, J. M., & Oliver, R. (2007). The relationship between language learning strategies, proficiency, age and self-efficacy beliefs: A study of language learners in Botswana. *System, 35*(3), 338–352. https://doi.org/10.1016/j.system.2007.01.003

Mahyudding, R., Elias, H., Cheong, L., Muhamad, M., Noordin, N., & Abdullah, M. (2006). The relationship between students' self-efficacy and their achievement. *Jurnal Pendidikdan Pendidikdan, Jil, 21*, 61–71.

Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Marzuki, M. (2014). Managing an effective English language laboratory in a polytechnic. *Studies in English Language and Education, 1*(2), 108–118.

Matthews, P. H. (2010). Factors influencing self-efficacy judgments of university students in foreign language tutoring. *The Modern Language Journal, 94*(4), 618–635. https://doi.org/10.1111/j.1540-4781.2010.01057

Mercer, S. (2011). Language learner self-concept: Complexity, continuity and change. *System, 39*(3), 335–346. https://doi.org/10.1016/j.system.2011.07.006

Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 275–294. https://doi.org/10.1111/j.1944-9720.2006.tb02266.x

Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning, 57*(3), 417–442. https://doi.org/10.1111/j.1467-9922.2007.00421.x

Morita, N. (2004). Negotiating participation and identity in second language academic communities. *TESOL Quarterly, 38*(4), 573–603. https://doi.org/10.2307/3588281

Nesse, R., & Lloyd, A. (1992). The evolution of psychodynamic mechanisms. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.

Nguyen, N., Shawa, J., Tyler, M., Pinkus, R., Best, C. (2015). Affective attitudes towards Asians influence perception of Asian-accented vowels. *International Phonetic Association*. https://www.internationalphoneticassociation.org/icphsproceedings/ICPhS2015/Papers/ICPHS0561.pdf. Accessed 10 Feb 2015.

Norton, B. (2000). *Identity and language learning: gender, ethnicity and educational change*. London: Longman.

Nowak, M., Plotkin, J., & Krakauer, D. (1999). The evolutionary language game. *Journal of Theoretical Biology, 200*, 147–162.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge, UK: Cambridge University Press.

Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 6*(4), 543–578.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing. *Reading and Writing Quarterly, 19*(2), 139–158. https://doi.org/10.1080/10573560308222

Parry, T., & Stansfield, C. W. (Eds.). (1990). *Language aptitude reconsidered*. Englewood Cliffs, NJ: Prentice Hall.

Pellegrino, V. A. (2005). *Study abroad and second language use: Constructing the self*. Cambridge, UK: Cambridge University Press.

Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research and applications*. Englewood Cliffs, NJ: Prentice Hall Merrill.

Podesva, R. (2006). *Phonetic detail in sociolinguistic variation: Its linguistic significance and role in the construction of social meaning*. Stanford, CA: Stanford University Press.
Rahimi, A., & Abedini, A. (2009). The interface between EFL learners' self-efficacy concerning listening comprehension and listening proficiency. *Novitas-Royal, 3*(1), 14–28.
Raoofi, S., Tan, B. H., & Chan, S. H. (2012). Self-efficacy in second/foreign language learning context. *English Language Teaching, 11*(5), 60–73.
Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., et al. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology, 2*, 1–12. https://doi.org/10.3389/fpsyg.2011.00271
Reiterer, S., Hu, X., Sumathi, T., & Singh, N. (2013). Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in Psychology, 4*, 1–13. https://doi.org/10.3389/fpsyg.2013.00782
Renninger, K. (2000). Individual interest and its implications for understanding intrinsic motivation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 373–404). San Diego, CA: Academic.
Rosenhan, D. (1973). On being sane in insane places. *Science, 179*, 250–258.
Rosenthal, R. (1985). From unconscious experimenter bias to teacher expectancy effects. In J. B. Dusek, V. C. Hall, & W. J. Meyer (Eds.), *Teacher expectancies* (pp. 37–66). Hillsdale, NJ: Erlbaum.
Rubio, F. (2007). Self-esteem and foreign language learning: An introduction. In F. Rubio (Ed.), *Self-esteem and foreign language learning* (pp. 2–12). Cambridge, UK: Cambridge University Press.
Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychology, 26*, 207–231.
Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modelling, goal setting. *Difficulties, 19*(2), 159–172. https://doi.org/10.1080/10573560308219
Schwarzer, R., & Jerusalem, M. (1993). *From measurement of perceived self-efficacy: Psychometric scales for cross-cultural research*. Berlin, Germany: Freie Universität.
Singleton, D. (2014). Apt to change: the problematic of language awareness and language aptitude in age-related research. *Studies in Second Language Learning and Teaching, 4*(3), 557–571. https://doi.org/10.14746/ssllt.2014.4.3
Stansfield, C. (1989). *Language aptitude reconsidered*. Washington, DC: ERIC Clearinghouse on Languages and Linguistics.
Sternberg, R. (2002). The theory of successful intelligence and its implications for language aptitude testing. In *Individual differences and instructed language learning.* Amsterdam: John Benjamins.
Tilfarlioğlu, F. T., & Ciftci, F. S. (2011). Supporting self-efficacy and learner autonomy in relation to academic success in EFL classrooms (a case study). *Theory and Practice in Language Studies, 1*(10), 1284–1294. https://doi.org/10.4304/tpls.1.10.1284-1294
Tracy, E., Bainter, S., & Satariano, N. (2015). Judgments of self-identified gay and heterosexual male speakers: Which phonemes are most salient in determining sexual orientation? *Journal of Phonetics, 52*, 13–25. https://doi.org/10.1016/j.wocn.2015.04.001
Williams, M., & Burden, R. (1997). *Psychology for language teachers*. Cambridge, UK: Cambridge University Press.
Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System, 39*(5), 510–522. https://doi.org/10.1016/j.system.2011.10.017
Woods, D. (2003). The social construction of beliefs in the language classroom. In P. Kalaja & A. M. F. Barcelos (Eds.), *Beliefs about SLA: New research approaches* (pp. 201–229). Dordrecht, The Netherlands: Kluwer Academic Publishers.
Yilmaz, C. (2010). The relationship between language learning strategies, gender, proficiency and self-efficacy beliefs: A study of ELT learners in Turkey. *Procedia Social and Behavioural, 2*(2), 682–687. https://doi.org/10.1016/j.sbspro.2010.03.084

Motivation and Personality in Language Aptitude

Nejra Rizvanović

Abstract Since individual differences (ID) have gained increasing popularity in second language acquisition (SLA) research, their five main areas of research, namely personality, aptitude, motivation, learning styles and learning strategies have crystallized. This section explores the relationship between two of these factors, personality and motivation, and foreign language aptitude (FLA). Previous research has shown a link between (1) openness and successful language attainment, (2) extraversion and higher fluency, and (3) empathy and pronunciation, among many others. In order to unveil possible relationships between FLA and personality and motivation, the LLAMA language aptitude testing battery and four questionnaires are used in this study: the Empathy Quotient by E.J. Lawrence, the Four Temperaments Test by Eric Jorgenson, the Big5 Personality Model, as well as a modified version of the BisBas personality scale which measures motivation. Results show that extrinsic motivation correlates negatively with LLAMA E and LLAMA compound scores, which suggests a superior status of intrinsic motivation in language acquisition. With regard to temperament, phlegmatics performed better in LLAMA B, E and D. Moreover, males had a higher mean than females in the compound LLAMA score, with a striking difference in LLAMA F (grammatical inferencing). The great amount of variety in scores reveals the significance of affective factors such as personality, motivation and empathy in the language learning process.

1 Introduction

This chapter delves into the complex nature of second language acquisition and aims to answer two questions, namely to what degree and in what way the psychological qualities of a person affect language learning. More specifically, the

N. Rizvanović (✉)
Department of Philosophy, MEi:CogSci Programme, University of Vienna, Vienna, Austria
e-mail: a01269367@unet.univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_6

affective domains of personality and motivation are investigated, and it is examined whether there is a relationship between certain types of characters, motivational levels and language aptitude. Hence, the aim of the study is trifold: (1) to uncover the relation between affective factors such as motivation, empathy and personality, (2) to shed light on the role of these affective factors in the language learning process, and (3) to clarify in which way they assist or hinder it.

1.1 ID Research

An influential figure in SLA research, Zoltan Dörnyei (2001a, 2001b, 2006), applies the ideas from the field of differential psychology to the SLA context, and stresses the importance of individuality in educational settings. The underlying assumption concerning ID research in linguistics is that people who differ in how they *are*, will also differ in how they learn a language. The five most significant facets which frame ID research include aspects related to learning (styles and strategies), personality, aptitude and motivation. Contrary to popular psychological practice, which studies individual subjects in relation to control groups, ID research seeks to understand the differences that exist in individuals.

According to Dörnyei (2001a, 2001b, 2006), IDs represent stable personal qualities shared by all individuals, but expressed to a different degree. Representing enduring features, IDs can account for the great degree of variance recorded in learners' success. As such, they also provide a reliable measuring tool that can predict the learning outcome. A case in point is a study conducted by De Raad (2000) that found IDs to be a stable predictor of the way humans think and behave. Similarly, studies by Sawyer and Ranta (2001), and Dörnyei and Skehan (2003) showed IDs to affect second language attainment, especially when a language is learnt through instruction.

The idea of language as a part of one's identity, and not as a mere means of communication, motivated a new way of looking at language. This time, the personality of the learner is considered to be a key factor in predicting success in language learning. Furthermore, the same thought leads researchers to conclude that the processes involved in learning a new language differ from other disciplines, as they are highly dependent on the identity of the individual. This idea has also been voiced by Dörnyei who proposed a model which places motivation and identity in the center of language learning (Dörynei, 2001a, 2001b; 2006; 2010).

Investigating IDs may impact numerous fields but in particular for education, it may have far-reaching consequences. For teachers it is of extreme importance to understand the differences that are present among students in order to provide sensitive instruction and achieve optimal results. Similarly, to properly study language and aptitude, it is imperative to pursue a more nuanced, rather than a categorical model.

1.2 Personality, Motivation and Empathy

1.2.1 Personality

Dörnyei (2006) describes personality as being the characteristic that differs most between human beings, i.e., the most individual one. It is because of this that it takes a central part in ID research. Hu and Reiterer (2009) refer to it as a stable, enduring formation of a person's being that influences their model of thought and behavior, and is therefore a reliable predictor of SLA success.

Motivation and anxiety have received a great deal of attention as influential factors that predict second language success. These factors are, however, considered to be cognitive rather than affective, as Dörnyei (2001a, 2001b, 2010) would argue. As a result, affective factors such as personality qualities have been set aside and rather neglected in research, as Biedroń (2011) points out. According to her, one of the reasons for this was the fact that affective factors have always been regarded as complementary to the more superior cognitive factors. She claims that another reason for the lack of investment in this field are the poor correlations found between personality dimensions and aptitude in previous research. She further explains that some researchers argue that affective and maturational factors do in fact have a more powerful influence than has been proved so far – a claim supported by the research study. The events that led to recognizing personality as a potentially powerful indicator of L2 attainment can be traced back to the fields of sociology and psychology. Psychologists argued that behavior is largely influenced by personality, and is a promising predictor of it. In addition to that, certain cross-cultural studies have shown the universal nature of traits which defy cultural boundaries. What is more, the popularity of ID research has elevated the priority of affective factors, as it defined personality as the most unique to all humans. These differences are advantageous for studying language attainment, as they assume different outcomes following different means. As such, it holds immense potential in L2 research.

With its rising popularity, the question arose as to which traits should be representative of personality, and various models were proposed, such as the Myers-Briggs Type Indicator. Researchers seek an exhaustive, yet selective representation of traits, which should be replicable and universal (Biedroń 2011). Eysenck and Eysenck (1976), and Tellegen (1982) proposed a three-trait model, Comrey (1970) one consisting of eight traits, Cattell et al. (1970) one of sixteen and McCrae and Costa (2003) suggested five higher-order traits. Applying the personality scales from the field of psychology, McCrae and Costa (2003) devised a personality model based on five main dimensions. They included openness to experience, conscientiousness, extraversion, agreeableness and neuroticism: Openness to experience is synonymous with curiosity and unconventionality in thinking and acting, as well as a general tendency to novelty. The five underlying traits associated with openness are fantasy, aesthetic pleasure, being open towards ideas, and values. Conscientiousness, on the other hand, is associated with being down to earth, ambitious, systematic and methodical. Qualities pertaining to this dimension are aspects like competence or self-discipline. What is more, it is a trait which is most likely to predict academic success and high levels of motivation. Extraversion implies activity, gregariousness, and that an individual has a tendency towards being externally stimulated. People with high levels of extraversion are the life of the party, loud and talkative, and always ready to engage. As Biedroń (2011) succinctly summarizes, the fundamental qualities of extraversion include warmth or an excitement-seeking character with very positive emotions. Agreeableness includes qualities of altruism, cooperative behavior, and generosity. The specific sub traits incorporate values like trust, concern for others or tender-mindedness. Finally, neuroticism is affiliated with emotional instability and lack of control in day-to-day situations, due to the tendency towards pessimistic and destructive feelings.

The final personality model which will be addressed in this section is Gray's BisBas scale (1981, 1982), which is short for behavior inhibition system and behavior activation system. It was first proposed in the 1970s, and argued for two underlying motivational systems which govern behavior and have a cognitive foundation. Gray found out that anxiety lies at the root of the inhibition system, and is responsive to punishment, non-reward, and novelty. It is biologically and evolutionally determined, as it prepares the individual for avoidance of undesirable outcomes. The activation system, in contrast, responds to reward and assumes a positive outlook. Individuals who lean towards the latter disposition tend to be more goal-oriented and motivated. The latter would therefore be more advantageous for language learning (Carver and White 1994; MacAndrew and Steele, 1991).

1.2.2 Motivation in SLA

The importance of motivation in all learning contexts, as well as its long-term importance in goal-achievement, is an undisputed fact that led to the acknowledgment of its reputation and as a result produced more interesting research in the field. This sub-section draws on this research with a special focus on a certain type of motivation, namely language learning motivation.

In order to take a deeper look into language learning motivation itself, we must turn to the theories of Mowrer, Lambert, Dörnyei, Ushioda and Gardner. According to Dörnyei (2006) and Dörnyei and Ushioda (2011), Robert Gardner is most known for his theory on integrative, as well as instrumental motivational orientation. The integrative motive can be considered an equivalent to intrinsic motivation, whereas instrumental motivation is synonymous with extrinsic motivation. The integrative motive voices the belief that the desire to achieve a goal originates within the learner, while the instrumental motive stems outside of the learner. In his book on the importance of individual differences in language learning, Skehan (1989) argues that the integrative motive is firmly grounded in the learner's personality, making it a more stable construct over time, whereas the instrumental orientation is much more dependent on outside stimuli, which are naturally unstable and situation-dependent. Integrative orientation seems to be the preferred choice when it comes to successful language learning, taking into account its nature to persist and endure over time.

What is the main line of differentiating between extrinsic and extrinsic orientation? What drives one into adopting one or the other motive, and can these overlap and evolve into the other, assuming the mutable nature of motivation? How do we trace this change? These are all questions that pose limitations on research of motivation, but also serve as basis for further initiatives in the field. Taking the limitations into account, the insights on motivation discussed so far can only serve as guidelines to help us decide which questions to ask and which aspects of motivation to focus on in our research.

1.2.3 Empathy

Dewaele and Wei (2012) argue that despite being an extensively researched concept, empathy is still difficult to describe due to its multifaceted nature. By definition, empathy is the ability to put yourself into the position of someone else including feelings and thoughts. But, as the definition implies, it includes multiple aspects, such as the emotional and the cognitive one. It comprises factors such as social self-confidence, even temperedness, sensitivity, non-conformity, or tolerance of ambiguity.

Guiora, Brannon and Dull (1972), as well as Rota and Reiterer (2009) have found high levels of empathy to correlate with pronunciation. An interesting clarification of this link comes from Guiora (1990). He traces it back to a psychological phenomenon known as ego flexibility. Guiora argues that the same can be translated to language learning, as there is a so-termed language ego, which functions in the same way. In this sense, higher permeability facilitates learning a new language, as the learner is more willing to adopt new views and is less defensive against outside stimuli. On the other hand, low permeability inhibits learning. Following this line of thought, Guiora also explains that children are much more successful in mastering a new tongue, since they adopt it while in the state of high ego permeability, which unfortunately decreases with time.

The current view on personality traits, as endorsed by personality psychologists, states that personality is rooted in our nature. In order to investigate possible links between personality traits (such as empathy) and aptitude, one has to take into account the social and cultural factors, and assume that these play a part in shaping the individual's language abilities. Focusing on this approach, Dewaele and Wei (2012) found cognitive empathy to correlate with gender, education level, and multilingualism, but not with bilingual background or foreign experience. Research which compared monolingual and bilingual children has shown that bilinguals scored significantly higher on tasks on executive functions which are involved in the processes of attention, selection, inhibition, shifting and flexibility. These are also referred to as higher cognition tasks, which attest to the existence of language influence on cognitive function. Bilinguals were also found to be more creative, and competent in abstract and symbolic representation. The benefits of knowing more languages on cognition have initially been accredited to having mastery over various language structures. Another explanation, however, has emerged, urging that it is actually the extensive cultural understanding which is the primary influence on cognition. It is clear at this point that the cognitive abilities mentioned so far relate closely to empathic ability. In this context, Guiora et al. (1972) advocate that empathy lies at the root of language learning, since learning requires an adoption of an unfamiliar identity. This requires a degree of openness and flexibility, as novel experiences need to be adopted and adjusted to the existing ones.

2 Hypotheses

As mentioned above, this study aims at investigating the following issues: the relation between affective factors such as motivation, empathy and personality, and the role of these affective factors in the language-learning process and in which way they assist or hinder it. It is argued that individual effort, motivation and perseverance affect the outcomes of language learning to a great extent, and that certain personalities will have areas of language in which they perform better or worse. This research is conducted with the aim of finding out which personality types excel or perform poorly in certain areas, and to measure to which extent the affective factors determine language success. The following concrete hypotheses are formulated:

H1: Higher levels of empathy are related to higher scores in LLAMA D (phonetic memory) and LLAMA E (sound-symbol correspondence).

H2: Higher conscientiousness levels are related to higher LLAMA F (grammar inferencing) and LLAMA B (vocabulary learning) scores.

3 Methodology

3.1 Participants

The sample consists of 19–35 year old participants ($N = 26$, males = 13, females = 13) who were invited to take part in this study. They were randomly selected for their academic background and language experience, but controlled for the criteria of age. The majority of participants was aged 22–25 years (M = 25.35), with the youngest participant being 19 and the oldest 35. Ten out of 26 participants were language students, all others were students of other degrees, and one was self-employed. Most were undergraduates or Master graduates and working, with one participant being in a PhD candidate. This places the majority of the participants in the tertiary education sector. As far as ethnic background is concerned, the countries of origin included Austria, Bosnia and Herzegovina, Bulgaria, Colombia, Croatia, Greece, Indonesia, Luxemburg, Pakistan, Slovenia and Syria. The number of spoken languages among the group ranged from two to four, with a mean of 2.72. The distribution of temperament types was heterogeneous, with sanguine ranking first (n = 12), followed by an equally distributed number of melancholic (n = 6) and phlegmatic (n = 6); choleric was represented the least (n = 2).

3.2 Instruments

To assess the language aptitude of the testing group, the participants partook in the LLAMA tests, developed by Paul Michael Meara of Swansea University (Meara, 2005). Due to the ethnic and linguistic diversity, as well as differing academic pursuits, the LLAMA tests provided a great basis for measuring aptitude as they use a pseudo-language for the assessment of skills. What is more, the language-free nature of the tests places each participant at the beginner's level, allowing for unwavering results. The tests are intended to measure four different language dimensions, with part B focusing on vocabulary learning, part D on sound recognition, part E on sound-symbol correspondence, and finally F on grammatical inferencing. Part B is based on visual stimuli, involving a set of 20 symbols which are assigned to arbitrary words. The level of aptitude is measured based on the number of correct matches the participants manage after the 120 s they have to memorize the items. Part D uses vocal stimuli based on a dialect of Northern Canada, where a speech engine generates ten sounds, one after the other. After hearing the sound string, participants are asked to identify target from non-target vocalizations, and their ability is assessed based on the number of correct identifications. Part E of the LLAMA aptitude tests relates to sound-symbol correspondence, where the participants have 120 s to learn the phonetic realizations of 24 phonemes. After the time has elapsed, they are presented with two combinations of two randomly conjoined phonemes, and have to choose the target from the non-target combination. Finally, part F consists of 20 grammatical constructions (i.e. distinct features for plural or gender) contained in 20 squares which participants are encouraged to open as many times as they wish in order to seek out important grammatical patterns. Furthermore, participants are instructed to note down any such patterns they identify. After 300 s, a symbol and two phrases appear on the screen. The task is to choose the target phrase to correspond to the symbol. Assessment follows on the basis of the number of correct matches.

Similar to the aptitude tests, the assessment of the personality dimensions centers around four questionnaires – the EQ by E.J. Lawrence (originally created by Baron-Cohen and Wheelwright, 2004), the 4Temp developed in 2014 by Eric Jorgenson and largely modelled after Eysenck (1967, 1973), the Big5 based on the work of Goldberg (1992), and a modified version of the BisBas personality scale. Notably, the Empathy Quotient (EQ) questionnaire measures four subgroups pertaining to the concept of empathy. They include cognitive empathy, emotional reactivity, social skills, and socially-desirable responses. The paper, therefore, takes the Lawrence's Empathy Quotient Questionnaire as a starting point, and sets to quantify the four components individually, but also measures the total empathy quotient by totaling the individual scores. The temperaments scale consists of 40 statements which participants rate on a five-point Likert scale (1=Disagree, 3=Neutral, and 5=Agree). The results are based on the representation strength of the type, calculated by averaging out the number of points each statement of a specific group (i.e. temperament type) received. Furthermore, the 50-item Big5 inventory measures five different personality dimensions: extraversion, agreeableness, conscientiousness, neuroticism and openness to experience.

Finally, the BigBas scale for motivational behavior was adapted to fit the constructs of intrinsic and extrinsic motivation. The shorter 17-item version was created for this purpose, and it contains five items pertaining to extrinsic, and four to extrinsic motivation component. Furthermore, it kept the original eight statements exploring the drive component of the BAS scale, as well as its fun-seeking component.

3.3 Procedures

All participants filled out a general questionnaire about their age, education, and language background (amongst others) before administering both the LLAMA tests and personality questionnaires. The participants received instructions face-to-face, via Skype, and via e-mail, and completed the tests either per Skype or in person. The time it took the participants to finish the tests ranged from 20 to 40 min.

4 Results

Data was analyzed using SPSS (Version 21) to test the following hypotheses: H1 assumed that higher levels of empathy are related to higher scores in LLAMA D (phonetic memory) and LLAMA E (sound-symbol correspondence). H2 predicted that higher conscientiousness levels are related to higher LLAMA F (grammar inferencing) and LLAMA B (vocabulary learning) scores. Both H1 and H2 failed to be supported as evident in the following analyses. However, light could be shed on the relation between affective factors such as motivation, empathy and personality, and the role of these affective factors in the language-learning process and in which way they assist or hinder it.

A Pearson's correlation was conducted between intrinsic motivation and LLAMA B, LLAMA D, LLAMA E, LLAMA F, and compound LLAMA scores, and between extrinsic motivation and the same LLAMA test scores. While there was no significant correlation found between intrinsic motivation and any of the LLAMA scores, there were two negative, moderate, significant correlations found between extrinsic motivation and LLAMA E ($r = -0.404$, $p = 0.041$), and extrinsic motivation and the total LLAMA score ($r = -0.402$, $p = 0.042$). These results suggest that external drives are not favorable in learning languages, and are especially ineffective in the sound-symbol correspondence task. This could be due to the fact that they lack the internal stability and continued interest to deal productively with language difficulties that arise in the process, which is necessary for motivational constancy and ultimate attainment of the language.

Furthermore, the sample was split in high and low performing groups based on the LLAMA B scores. Pearson's correlation tests were then run to assess the correlation that neuroticism has with intrinsic motivation, conscientiousness, and openness for each level of LLAMA B performance. There were significant correlations found within the high-performing group of the LLAMA B, where the neuroticism score correlated strongly and negatively with intrinsic motivation ($N = 13$, $r = -.684^{**}$, $p = .010$), conscientiousness ($N = 13$, $r = -.627^{*}$, $p = .022$), and openness ($N = 13$, $r = -.623^{*}$, $p = .023$). To the contrary, there were no statistically significant associations within the low scoring LLAMA B group. The results suggest that, when it comes to vocabulary learning, the higher the neuroticism score is, the less intrinsic motivation is present, and the less open or conscientious an individual is, but only for those who score high on the LLAMA B test (Table 1).

The same Pearson's correlations were conducted with high and low performing groups with respect to the LLAMA F. A statistically significant correlation surfaced within the low-performing group, where neuroticism was negatively associated with intrinsic motivation ($N = 12$, $r = -.767^{**}$, $p = .004$) and openness ($N = 12$, $r = -.719^{**}$, $p = .008$). With regard to grammatical inferencing, results suggest that neuroticism has a stronger, unfavorable influence on the low-performing group than the high-performing group, as it negatively affects intrinsic motivation and openness, which are believed to underlie successful language attainment (Table 2).

Table 1 Correlation matrix for high and low groups in Llama B and the affective factors

LLAMA_B			Motiv_intrins	C_score	O_score
High_group	N_score	Pearson correlation	−684**	−627*	−623*
		Sig. (2-tailed)	.010	.022	.023
		N	13	13	13
Low_group	N_score	Pearson correlation	.208	.383	.327
		Sig. (2-tailed)	.495	.196	.275
		N	13	13	13

Table 2 Correlation matrix for high and low groups in Llama F and the affective factors

LLAMA_F			Motiv_intrins	O_score
High_group	N_score	Pearson correlation	.085	.347
		Sig. (2-tailed)	.774	.225
		N	14	14
Low_group	N_score	Pearson correlation	−.767**	−.719**
		Sig. (2-tailed)	.004	.008
		N	12	12

Table 3 Group statistics for temperament types and aptitude scores

	Temperament_type	N	Mean	Std. deviation	Std. error mean
LLama_B	Melancholic	6	44.17	20.595	8.408
	Sanguine	12	53.33	20.487	5.914
	Choleric	2	50.00	21.213	15.000
	Phlegmatic	6	55.00	21.679	8.851
LLama_D	Melancholic	6	34.17	17.151	7.002
	Sanguine	12	32.50	17.255	4.981
	Choleric	2	27.50	10.607	7.500
	Phlegmatic	6	38.33	15.384	6.280
LLama_E	Melancholic	6	71.67	19.408	7.923
	Sanguine	12	75.83	21.515	6.211
	Choleric	2	50.00	70.711	50.000
	Phlegmatic	6	78.33	13.292	5.426
LLama_F	Melancholic	6	53.33	34.448	14.063
	Sanguine	12	46.67	31.431	9.073
	Choleric	2	55.00	7.071	5.000
	Phlegmatic	6	51.67	30.605	12.494
Llama_totale	Melancholic	6	203.3333	58.87841	24.03701
	Sanguine	12	208.3333	66.54914	19.21108
	Choleric	2	182.5000	88.38835	62.50000
	Phlegmatic	6	223.3333	54.00617	22.04793

To understand whether scores in the aptitude tests differed based on the temperament types, one-way ANOVAs were performed. These analyses showed that the effect of temperament types on Llama B scores was insignificant $F(3,22) = .339$, $p = .798$. Similarly, there was no statistical difference in the mean scores for temperament and Llama D scores $F(3,22) = .272$, $p = .845$; temperament and Llama E scores $F(3,22) = .759$, $p = .529$; temperament types and Llama F scores $F(3,22) = .091$, $p = .964$, and the compound Llama score $F(3,22) = .235$, $p = .871$.

For illustration purposes, Table 3 below displays the different temperament types and their performance on the aptitude tests. When looking at the means, the phlegmatic temperament outperformed the other three in three out of four aptitude tests, namely in the LLAMA B, the LLAMA D, and the LLAMA E.

Next, a Pearson Chi-square test was administered to test the strength of association between high and low aptitude groups and temperament. With respect to Llama B, the test did not yield a statistically significant association, χ^2 (3, $N = 26$) = 1.000, $p = .801$. There was no statistically significant association between temperament and the Llama D scores for the high and low groups, respectively, χ^2 (3, $N = 26$) = 2.058, $p = .560$, nor the Llama E scores and temperament, χ^2 (3, $N = 26$) = .394, $p = .941$. No significant association was found between temperament and the llama F scores either, χ^2 (3, $N = 26$) = .516, $p = .915$.

Additionally, a multiple linear regression was run to predict the individual aptitude scores from the affective factors. All variables were tested for linearity with q-q and scatter plots, which all showed a linear distribution. Results show that BAS drive, BAS fun, intrinsic and extrinsic motivation, the OCEAN scores, temperament type, and the four different facets of empathy did not significantly predict LLAMA B scores, $F(14, 11) = .567$, $p = .843$, $R^2 = .419$. None of the fourteen variables added significantly to the prediction, $p = >.2$

Similarly, the affective factors did not significantly predict LLAMA D scores, $F(14, 11) = .519$, $p = .867$, $R^2 = .398$, $p = >.2$, nor the LLAMA F scores, $F(14, 11) = .838$, $p = .629$, $R^2 = .516$, $p > .1$.

It is worth noting that the multiple regression analysis run for the Llama E scores showed affective factors to indeed significantly predict it, $F(14, 11) = 3.2$, $p = .029$, $R^2 = .804$. More specifically, four of the fourteen variables – intrinsic motivation, neuroticism score, temperament type, and the emotional reactivity facet of empathy – added significantly to the prediction ($p < .05$) with cognitive empathy being marginally significant ($p = .055$). The rest of the variables, however, did not add significantly to the variance in the LLAMA E scores, $p > .1$.

Overall, the results indicate an important role of affective factors in language aptitude. More accurately, they show that affective factors such as temperament, neuroticism and empathy can reliably predict performance in the sound-symbol correspondence task.

Moreover, since the affective factor of empathy is often influenced by gender, two independent samples t-tests were conducted for differences between males and females in empathy scores. The results show that there is a significant difference between females ($M = 4.46$, $SD = .79$) and males ($M = 3.73$, $SD = .84$) in overall empathy; $t(24) = -2.3$, $p = 0.03$. Similarly, a significant difference was found between females ($M = 1.15$, $SD = .31$) and males ($M = 0.79$, $SD = .37$) for social empathy; $t(24) = -2.76$, $p = 0.011$.

Finally, Pearson's correlations between openness and intrinsic motivation, and openness and cognitive empathy were conducted. The results suggest that there is a particularly strong bond between openness and intrinsic motivation ($N = 26$, $r = .530^{**}$, $p < .05$), and between openness and cognitive empathy ($N = 26$, $r = .497^{**}$, $p < .05$). This implies that an individual who is open to novel experience and has an appreciation for new cultures (as stated by definition) might more easily take on another's perspective or mental state (cognitive empathy), which can have a positive influence on the intrinsic motivation.

5 Discussion

The purpose of the study was to examine the impact of affective factors on language learning, with the aim of highlighting the importance of malleable factors and learned behaviors to inherited ones. This study also aimed to cast light on areas of personality, empathy and motivation with regard to their effect on personality.

Albeit it was found that affective factors do play a role in language learning to a certain extent, H1 and H2 failed to be supported due to statistically insignificant results.

H1, which assumed that higher levels of empathy are related to higher scores in LLAMA D (phonetic memory) and LLAMA E (sound-symbol correspondence), could not be supported as participants with the highest empathy did neither show higher scores in the phonetic memory, nor in the sound-symbol correspondence task. These results were surprising due to other research in the field which does demonstrate that empathy has a significant influence on language abilities (Guiora et al., 1972; Rota and Reiterer, 2009). A reason for the outcome of this study might lie in the research design, since more participants might have led to a different outcome. What is interesting, however, is that females have a higher overall and social empathy, which supports the findings by Dewaele and Wei (2012).

Likewise, the analyses could not support H2 which predicted that higher conscientiousness levels are related to higher LLAMA F (grammar inferencing) and LLAMA B (vocabulary learning) scores. This could suggest that there are different mechanisms involved in vocabulary and grammar acquisition which do not necessarily require conscientious and thorough work. Even though previous research suggests conscientiousness to be linked with higher motivation and academic success (Biedroń, 2011), it is still unknown what type of motivation (intrinsic or extrinsic) is at play, and where the line of division lies. Moreover, Jilka (2009) suggests additional, equally important mechanisms which he found to contribute to proficiency in this area, such as practice and experience. These mechanisms were considered by grouping scores of the vocabulary learning task into high and low. It has been shown that neuroticism was, indeed, a negative influence on intrinsic motivation, conscientiousness, and openness. Neuroticism was also shown to correlate negatively with intrinsic motivation and openness in the low group, with regard to the grammatical inferencing scores. Different results within these groups and across different aptitude tests suggest that affective factors play a moderating role in explaining differences in language success.

With regard to the motivation factor, interesting results surfaced. They include a highly significant correlation between intrinsic motivation and openness to experience, as well as between intrinsic motivation and cognitive empathy. Since extrinsic motivation was seen to affect the learning outcome negatively (see Table 2), it is safe to assume that intrinsic motivation is advantageous to this purpose. Equally important in this scenario are cognitive empathy and openness, as factors which positively influence or underlie intrinsic capability. The findings of this study resonate with the belief that empathy and openness, which is its necessary component, underlie successful L2 attainment (Guiora et al., 1972; Rota and Reiterer, 2009). In the light of these findings, one can say that cognitive empathy, openness and intrinsic motivation are beneficial components in language learners. Nevertheless, it is not yet clear whether the constellation of these skills affect certain language areas (i.e. pronunciation), or contribute to a higher aptitude overall. What is more, the findings of this study agree with the ideas of Skehan (1989) and Gardner (Dörnyei, 2006; Dörnyei and Ushioda, 2011), who claim superiority of intrinsic over extrinsic motivation in language learning, due to the fact that the former is viewed as more stable since it originates within, and thus forms part of a learner's personality. It has also been claimed that intrinsic motives are especially advantageous because of their enduring quality, as they are a natural push for the learner when he or she faces challenges in the language learning process. Since intrinsic motivation is an end to itself, and not purely a mean, it is strongly attached to the goals of the learner, and thus generates motivation needed to reach that goal (Dörnyei, 2006; Dörnyei and Ushioda, 2011).

What is more, the statistical analysis did not yield significant results among the variables of temperament types and LLAMA aptitude tests. The results, however, allude to a superior status of phlegmatics over other temperaments in language learning, as they outperformed the other three in three out of five aptitude tests (LLAMA B, E, D; see Table 3) and hold the highest compound score ($M = 223.333$).

Furthermore, a multiple regression analysis revealed the significance of intrinsic motivation, neuroticism, temperament and emotional reactivity in predicting LLAMA E scores. Since phlegmatic types are characterized as peace keepers and conflict mediators, and are least competitive of all temperaments, it can be assumed that such a constellation of traits would be beneficial to language learning.

6 Conclusion

In this study the hypothesis that higher levels of empathy are related to higher phonetic memory and sound-symbol correspondence could not be supported; the same holds for the assumption that higher conscientiousness levels are related to higher grammar inferencing and vocabulary learning skills. Yet, results clearly indicate that affective factors do play a role in language learning to a certain extent. When investigating the findings of this study it needs to be considered that the low number of participants might account for the non-significant results regarding the research hypotheses, as well as for the uneven distributions among certain categories (i.e., temperament type). However, even with enough participants this study would have its limitations.

For instance, with respect to the affective factors, it is necessary to investigate them together with the cognitive aspects in order to infer proper conclusions on the matter. Further important aspects to consider are learner's strategies and self-image *(self-efficacy),* and none the least the teacher's role in the learning process, which all affect the motivation necessary for goal-achievement and the consequent language success.

Additionally, when studying the influence of personality factors on language skills, data on mechanisms which do not necessarily require conscientious work, such as experience, should be collected as a complement to the existing conscientiousness data (see Jilka, 2009).

Furthermore, the multiple regression analysis which yielded significant results regarding intrinsic motivation, neuroticism, temperament and emotional reactivity predicting LLAMA E scores needs to be viewed critically. It would be beneficial to replicate this measure with an equal number of participants with regard to temperament types with the aim of generating more reliable data.

When it comes to temperament types the question of to what extent they are classifiable into one specific category arises, since the results often showed participants having a nearly equal combination of two temperament types (e.g. melancholic and phlegmatic). In order to achieve more insight into the influence of temperament on aptitude, it would be beneficial to observe temperament on a spectrum rather than categorically, and restructure the questionnaire to fit those needs.

In general, further investigation with an equal and greater number of participants, and homogeneous groups could procure more substantial evidence and thus reinforce the belief that personality dimensions and other affective factors are reliable predictors of SLA outcomes.

To conclude, it needs to be noted that one can assume a fairly indispensable role of affective factors in SLA processes, as these should equal the importance of cognitive aspects in predicting language learning outcomes. Dörnyei (2001a, 2001b, 2006) and De Raad (2000) have found personality traits to be stable predictors of language success, as they represent the most unique aspect of each human being. Still, a part of the problem in acknowledging the importance of affective factors in language aptitude seems to be their categorization as cognitive factors (Dörnyei 2001a, 2001b, 2010). Nevertheless, it would be impractical and unreasonable to ignore these aspects when studying language aptitude, especially considering the importance of personality as a drive for behavior, way of thinking, and, ultimately, a predictor of how we learn.

Furthermore, disregarding the differences that exist among learners of foreign languages and assigning them to crude categories shuns the potential of ID to provide sensitive and unique instruction which could train language skills effectively. Dörnyei (2001a, 2001b, 2006) argues throughout his research that, even though the degree to which personality traits in an individual are expressed varies, they nevertheless represent enduring and stable features, which hold enormous potential in accounting for variance in language success. And, as such, they should be given credit and studied accordingly. Moreover, their universal nature makes them an indispensable component in studying language. Finally, these differences are advantageous for studying language attainment, as they assume different outcomes following different means, thus holding immense potential for studying L2 attainment.

As highlighted throughout this section, language should be seen as a faculty composed and influenced by various components – biological, cognitive, personal, cultural and historical. Affective factors such as empathy, motivation, and personality cannot be disregarded when investigating language acquisition and aptitude. It is, to this purpose, necessary to include all these factors when drawing conclusions from observable data, as their omission could lead to distorted interpretations.

References

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*(2), 163–175.

Biedroń, A. (2011). Personality factors as predictors of foreign language aptitude. *Studies in Second Language Learning and Teaching, 1*(4), 467–489. https://doi.org/10.14746/ssllt.2011.1.4.2

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *The handbook for the sixteen personality factor questionnaire*. Champaign, IL: Institute for Personality and Ability Testing.

Comrey, A. L. (1970). *Comrey personality scales*. San Diego, CA: Educational and Industrial Testing Services.

De Raad, B. (2000). *The big five personality factory. The psycholexical approach to personality*. Hogrefe & Huber: Berlin, Germany.

Dewaele, J. M., & Wei, L. (2012). Multilingualism, empathy and multicompetence. *International Journal of Multilingualism, 9*(4), 352–366. https://doi.org/10.1080/14790718.2012.714380

Dörnyei, Z. (2001a). *Teaching and researching motivation*. Harlow, UK: Longman.

Dörnyei, Z. (2001b). *Motivational strategies in the language classroom*. Cambridge, UK: Cambridge University Press.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*(1), 42–68.

Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247–267). London: Continuum.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Harlow, UK: Pearson Education.

Eysenck, H. J. (1967). *The biological basis of personality. Springfield, III: Thomas. Open four temperaments scales [Psychometric test]*. Retrieved from http://personality-testing.info/tests/O4TS/

Eysenck, H. J. (1973). *Eysenck on extraversion*. London: Crosby Lockwood Staples.

Eysenck, H. J., & Eysenck, S. B. G. (1976). *Psychoticism as a dimension of personality*. London: Hodder and Stoughton.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26–42 Big-five factor markers [Psychometric test]. Retrieved from https://openpsychometrics.org/tests/IPIP-BFFM/.

Gray, J. A. (1981). A critique of Eysenck's theory of personality. In H. J. Eysenck's (Ed.), *A model of personality* (pp. 246–276). Berlin, Germany: Springer Verlag.

Gray, J. A. (1982). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. New York: Oxford University Press.

Guiora, A. Z. (1990). A psychological theory of second language pronunciation. *Toegepaste Taalwetenschap in Artikelen [Journal of Applied Linguistics], 41*(1), 5–14.

Guiora, A. Z., Brannon, R. C., & Dull, C. Y. (1972). Empathy and second language learning. *Language Learning, 22*(1), 111–130.

Hu, X., & Reiterer, S. (2009). Personality and pronunciation talent in second language acquisition. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 97–130). Berlin, Germany: Mouton de Gruyter.

Jilka, M. (2009). Talent and proficiency in language. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 1–16). Berlin, Germany: Mouton de Gruyter.

MacAndrew, C., & Steele, T. (1991). Gray's behavioral inhibition system: A psychometric examination. *Personality and Individual Differences, 12*(2), 157–171.

McCrae, R. R., & Costa Jr., P. T. (2003). *Personality in adulthood: A five-factor theory perspective* (2nd ed.). New York: Guilford Press.

Meara, P. (2005). *LLAMA aptitude test series directory*. Retrieved from http://www.lognostics.co.uk/tools/llama/

Rota, G., & Reiterer, S. (2009). Cognitive aspects of pronunciation talent: How empathy, mental flexibility, working memory and intelligence interact with phonetic talent. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 67–96). Berlin, Germany: Mouton de Gruyter.

Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and instructed second language learning* (pp. 319–353). Oxford, UK: Oxford University Press.

Skehan, P. (1989). *Individual differences in language learning*. New York: Routledge, Chapman and Hall.

Tellegen, A. (1982). *Brief manual for the multidimensional personality questionnaire*. Unpublished manuscript. Minneapolis, MN: University of Minnesota.

Part III
Language Aptitude in Relation to Neuroscience and Musicality

The Neuroanatomical Correlates of Foreign Language Aptitude

Sabrina Turker, Susanne M. Reiterer, Peter Schneider, and Annemarie Seither-Preisler

Abstract In this chapter language aptitude will be looked at from a neuroscientific perspective and it will be discussed to which extent foreign language aptitude, i.e., foreign language learning potential or ability, is influenced by brain morphology, working memory and musical ability. The first part hence serves as a theoretical introduction and brief narrative review on past research, whereas the second part deals with a study on the role of the morphology of auditory cortex in German-speaking individuals with high and low language aptitude. In this study, MRI scans of German monolingual native speakers (N = 30; aged: 20–40 years) were analyzed and the auditory cortices of the participants with particularly high and those with particularly low language aptitude were compared. On the behavioral level, significant correlations could be found between speech imitation aptitude, English pronunciation skills, musicality and language aptitude as measured by the Modern Language Aptitude Test (MLAT). Especially the number of instruments played and working memory capacity showed significant correlations with aptitude measures. Moreover, it became clear that adults with very high language aptitude scores had more complete posterior duplications of Heschl's gyrus in the right hemisphere and thus a differently developed primary auditory cortex. The results are in accordance with what research has previously shown in musically gifted children and could

S. Turker (✉) · S. M. Reiterer
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Wien, Austria
e-mail: susanne.reiterer@univie.ac.at

P. Schneider
Department of Neurology, Division of Neuroradiology and Section of Biomagnetism, University of Heidelberg, Heidelberg, Germany

A. Seither-Preisler
Centre for Systematic Musicology, Karl-Franzens University Graz & BioTechMed, Graz, Austria
e-mail: annemarie.seither-preisler@uni-graz.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_7

reignite the discussion of the importance of right-hemispheric brain areas for language processing.

1 Introduction

People have always been fascinated by the simple fact that some individuals are just better at doing something, e.g., playing an instrument, singing or learning a language. The most striking aspect of this being better though, is that those seemingly more gifted individuals mostly do not put much effort into acquiring the skills and do need very little time to reach a high achievement or proficiency level. Such individuals are considered to possess an innate potential for achieving high ability in a certain domain, i.e., they have a certain aptitude for something (Al-Shabatat, 2013; Gagné, 1995, 2002, 2005; Nardo & Reiterer, 2009; Stern & Neubauer, 2013). How can these obvious differences in learning potential be explained? Simply stating that just motivation or just a genetic predisposition are the only obvious reasons for these differences in achievements would not do justice to this complex issue. A large variety of aspects, including environmental factors (e.g., educational background, family), personality factors (e.g., motivation, intra-/extraversion) (Biedroń, 2011a, 2011b, 2012; Dörnyei, 1998, 2006; Dörnyei & Ryan, 2015) and biological factors (e.g., working memory capacity) (Wen, 2012, 2016; Wen, Skehan, & Biedroń 2017) have a considerable impact on language aptitude (Biedroń & Pawlak, 2016a, 2016b; Brown, 2006; Carroll, 1990; Ganschow & Sparks, 1995; Granena & Long, 2013; Li, 2015, 2016; Singleton, 2017; Sparks & Ganschow, 2001; Sparks, Humbach, Patton, & Ganschow, 2011). Also musical abilities, such as playing an instrument or singing, may influence language aptitude and language proficiency (Christiner & Reiterer, 2013; Nardo & Reiterer, 2009; Rota & Reiterer, 2009). What is of central importance, however, is that the genetic component of aptitude cannot be denied and has been shown to be one of the most dominant as far as linguistic or musical giftedness are concerned (Seither-Preisler, Parncutt, & Schneider, 2014; Serrallach et al., 2016).

This chapter deals with the interconnectedness between a number of cognitive factors, namely language aptitude or henceforth also language learning ability, working memory, musicality, and linguistic and musical background (experience). Additionally, neuroanatomic aspects shall be taken into consideration as well and a focus will therein be paid to the structure of Heschl's gyrus (henceforth HG), the region responsible for auditory processing.

The following research questions shall be answered through the results achieved in this research project:

1. To what extent does language aptitude depend on other cognitive variables such as musicality and working memory?
2. Can the structure of the primary auditory cortex (more specifically Heschl's gyrus) be seen as a neuroanatomical marker of language aptitude?

The behavioral measurements of the research study included language aptitude scores as assessed by the Modern Language Aptitude Test (Carroll & Sapon, 1957;

henceforth MLAT), English pronunciation scores, Hindi speech imitation scores and working memory skills assessed through digit span, non-word span and backward span. The participants of the study went through behavioral assessments, and MRI scans were taken for auditory cortex segmentation.

1.1 Language Aptitude: Clarifying the Concept

The concept of language aptitude has been dealt with in the introductory chapter and shall only be revised briefly in this chapter.

For over 50 years after the birth of the discipline, language aptitude was defined as an outstanding ability facilitating foreign language learning in terms of that an individual acquires a foreign language very quickly and with very little effort (Carroll, 1958, 1962, 1973, 1990; Stansfield & Reed, 2004). More recent approaches (Robinson, 2005) accurately describe language aptitude as a particular strength in cognitive abilities, which are especially drawn upon during foreign language learning. In the past years, researchers have also shifted their foci more towards investigating the influence and importance of individual differences between language learners (Biedroń, 2015; Dörnyei & Ryan, 2015; Dörnyei & Skehan, 2003; Dörnyei, 1998, 2006; Robinson, 2002, 2012; Skehan, 1986, 2002; Spolsky, 1995; Wen et al., 2017). With regard to the components of language aptitude, some theoretical advancements have been made and it has been questioned if the various components of language aptitude might be relevant for different levels and contexts of learning (Abrahamsson & Hyltenstam, 2008; Artieda & Muñoz, 2016).

Certainly, language aptitude is only one of the many factors accounting for the individual differences found in SLA (second language acquisition) research, but it certainly plays a vital role. Figure 1 is a recent model that was developed by Seither-Preisler et al. (2014) and even though it was developed based on musical ability, it can be applied to linguistic contexts as well. As displayed in the model, an individual's aptitude profile has an impact on maturational plasticity resulting in a certain competence profile. Additionally, aptitude is a potential that leads to intrinsic motivation, which consequently leads to good practicing behavior and thus learning-induced plasticity. Last but not least, pedagogical intervention may inhibit or support motivation and practicing behavior and cause differences in the overall aptitude profile. On the right side, the neuroscientific aspects of aptitude are given, described here as pre-existing anatomical factors, i.e., brain morphology, and higher neural efficiency. Figure 1 helps us acquire a better understanding of the complex nature of the construct of aptitude and the many components that are essential for its development. It sees aptitude as the innate capacity to develop a certain talent.[1]

[1] Different models of giftedness exist and cannot be dealt with explicitly in this chapter. For a complete overview on distinctions between giftedness and talent, please refer to Gagné (2002, 2005), and for a summary on approaches towards these concepts refer to Al-Shabatat (2013).

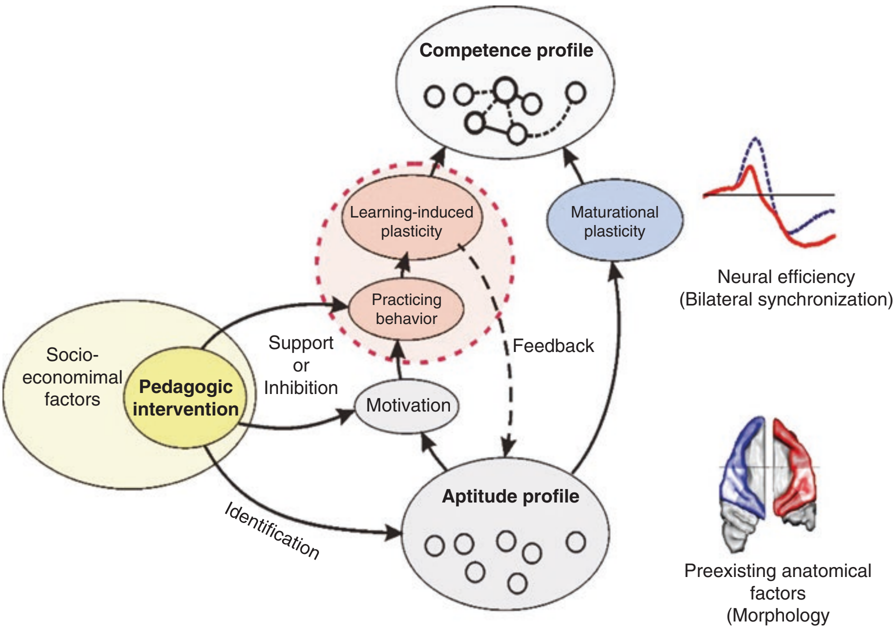


Fig. 1 A model of the various factors influencing (language) aptitude by Seither-Preisler et al. (2014)

1.2 *Language Aptitude from a Neuroscientific Perspective*

While language has been one of the hot topics in neuroscientific research, language aptitude or language learning ability has gained very little attention until recently. There is no doubt that the cognitive neurosciences offer a great opportunity to gain a better insight into the internal workings of the brains of very gifted individuals, too (Biedroń, 2015; Biedroń & Pawlak, 2016a, 2016b).

With respect to aptitude research, the three sources of ability differentiation related to the brain have been claimed to be (1) neural conduction velocity, (2) neural efficiency, and (3) grey and white matter volumes (Biedroń, 2015). In simpler terms, language aptitude can be looked at from different viewpoints: the connectivity of brain regions, the speed of signal transmission and the structure of the brain. Structural variations in brain morphology include differences in gyral and sulcal patterns (Schumann, 2004). Typically, the same major gyri and sulci are found in every single brain but it is their exact shape and location that may vary substantially (Kemmerer, 2015). But what do these differences tell us? So far, it remains quite a debate why differences in brain function or anatomy exist, where they come from and in how far they affect behavioral output. Structural variation could be innate, early intrauterine-determined or developed during infancy or childhood (Schumann, 2004). Schumann (2004) summarizes the aforementioned possibilities by stating that there are five sources of variation among brains, namely genetics, development, experience, degeneracy and individual appraisal systems. Long-term studies, such as Seither-Preisle et al. (2014) and Serrallach et al. (2016), investigated the auditory cortices of children at different points in time and have shown that gyral and sulcal patterns seem to be relatively stable and do not change over long periods of time (8 years in their study so far). Other studies on the macrostructure of auditory cortex also point towards a strong genetic component (Chi, Dooling, & Gilles, 1977; Hulshoff et al., 2006; Leroy et al., 2015).

Neuroanatomical differences can be found in every part of the brain but recently, the auditory cortex and in particular the shape of HG has gained increasing attention as it is known to be essential for many language- and music-related processes. As a matter of fact, thanks to the auditory cortex, we, focusing on the linguistic importance of it, perceive "the sounds of human speech […], complex acoustic patterns, precisely sculpted by a set of independently adjustable articulatory organs" (Kemmerer, 2015, p. 111). As the name already reveals, the auditory cortex is the location of primary and secondary auditory processing in the human brain. It extracts information from speech and non-speech stimuli and passes them on to other areas (for details, see Kemmerer, 2015). Concerning the gross anatomy of auditory cortex, it can be said that it occupies a great part of the superior temporal gyrus (STG), as well as a lower bank of the lateral sulcus in the temporal lobe (Benner et al., 2017; Hackett, 2009, 2015; Kemmerer, 2015).

The primary auditory cortex mostly comprises the postero-medial part of HG. Most humans possess a single or paired HG, which is often also termed posterior duplication or bifid HG. HGs can be divided by a sulcus and the core thus occupies both gyri (Benner et al., 2017). The primary auditory cortex also possesses a so-called tonotopic organization and responds to the frequencies as processed in the cochlea (Bear, Connors, & Paradiso, 2007; Hackett, 2009; Purves et al., 2001).

What makes HG so interesting for research is not only its importance for linguistic and music processing, but also the fact that there seems to be considerable structural variation between individuals. One difficulty when investigating HG in structural scans, however, is the lack in agreement as to where exactly HG begins and ends. The research study presented in this chapter relies on anatomical marks used by Peter Schneider and colleagues in various projects (Benner et al., 2017; Schneider et al., 2002; Schneider, Sluming, Roberts, Bleeck, & Rupp 2005; Seither-Preisler et al., 2014; Serrallach et al., 2016). For an excellent discussion on various approaches on MRI analysis of HG, see Abdul-Kareem and Sluming (2008).

1.3 Investigating Neuroanatomic Differences Due to Aptitude

After the first studies focusing on the neurological basis underlying this abstract concept of language aptitude, it was claimed that linguistic talent is a result of greater neurocognitive flexibility and bilateral processing, which enables individuals to acquire a language fast and effortlessly in comparison to age-matched peers (Biedroń, 2015; Schneiderman & Desmarais, 1988a, 1988b). In simpler terms, linguistic ability was thought to depend on the fact that both hemispheres work together more successfully and that the brain is more apt to adapt to learning, which could to some extent be confirmed in previous studies. This approach emphasizes the functional part of language aptitude but structural variation may also account for the variability in language learning ability encountered in individuals. Doubtlessly, language aptitude could be the result of "inborn functional and structural/anatomical characteristics as well as an individual brain response to an idiosyncratic experience of learning a foreign language" (Biedroń, 2015, p. 16).

There has been a considerable amount of neuroscientific research on atypical learning and brain-related issues, foreign language learning and language processing more generally.[2] Research on foreign language aptitude, though, has only received very little attention (functionally and structurally) and very few studies have focused on the possible structural differences encountered in individuals with differing degrees of linguistic ability or aptitude. Let us begin with a short review on musical ability, working memory and then language aptitude.

One of the core factors of this chapter is the relationship between working memory capacity (for details, refer to Baddeley, 2003a, 2003b; Baddeley & Hitch, 1974, 2000) and language aptitude, which has gained momentum in the past years (Miyake & Friedman, 1998; Sawyer & Ranta, 2001; Wen, 2012, 2016; Wen & Skehan, 2011; Wen et al., 2017). Studies comparing the two abilities have definitely confirmed the impact of the latter on numerous language-related skills, e.g. faster and highly successful first and foreign language learning (Ellis & Sinclair, 1996; Kormos & Sáfár, 2008; Linck et al., 2013; Sáfár & Kormos, 2008). In other words, individuals who have significantly better working memory capacity seem to be much more successful when it comes to foreign language learning (Biedroń, 2012; Van den Noort, Bosch, & Hugdahl, 2006). Just highlighting general working memory capacity and its significance for foreign language learning ability is not enough though. There are differences between specific working memory components, how they can be tested and in how far they are relevant to and predictive for the known components of foreign language learning ability (Baddeley, 2003a, 2003b, 2017). Although many researchers have insisted on the concept of working memory being almost equivalent to language aptitude, i.e., language learning ability could be replaced by working memory capacity, other studies have questioned this theory (Winke, 2013).

Musical ability/musicality/musical aptitude[3] and also its relationship to linguistic ability have been addressed in a variety of studies. As a matter of fact, language and music are two auditory phenomena that share numerous similarities. They are both structural systems, they are conveyed by sounds and they are specific to humans (i.e., universals). Moreover, both rely on intentionality, require a theory of mind and

[2] For a complete overview on language processing more generally, see excellent reviews by Price (2010, 2012) and Richardson and Price (2009).

[3] All three terms have been used here to avoid possible limitations due to the usage of one of the many terms suitable here. They certainly go hand in hand and going into detail about terminological issues on music ability research would go far beyond the scope of this chapter. Any kind of musical ability, talent or aptitude is included in this section.

are acquired by children without much effort (Besson & Schön, 2001). Given the common features of music and language, the suggestion of similar underlying brain functions and hence a common network in the brain seems logical. Several studies have already confirmed the overlapping auditory processing mechanisms involved in both language and music processing and the possible benefits of speech processing from the more demanding processing of music (Koelsch, 2005; Patel, 2011). Also, it has been postulated that either the left or right auditory cortex may be of greater importance for the one or the other due to the hypothesis that the left and right auditory cortex may be responsible for processing auditory input of different nature, the left essential for language, the right for music (Serrallach et al., 2016; Zatorre, Belin, & Penhune, 2002).

Music training does not only have a positive effect on language processing and the development of linguistic skills; other cognitive abilities, such as reading, attention and memory, seem to benefit from it (Besson & Schön, 2001; Koelsch, 2005; Milovanov & Tervaniemi, 2011; Patel, 2011; Seither-Preisler et al., 2014; Serrallach et al., 2016). A considerable amount of research has proven that music training has positive, long-lasting biological benefits on auditory functioning and may also lead to neuroanatomic differences such as morphological changes in the precentral gyrus, the motor brain areas or Heschl's gyrus. But these benefits extend far beyond simple enhancements regarding perceptual abilities and have been shown to positively impact non-auditory functioning involving working memory and intelligence, i.e., high-order aspects of cognition. It is argued that auditory training changes the processing of sound stimuli in the brain and therefore specific life experiences may cause specific functional changes. As a matter of fact, musicians' perceptual, language and high-level cognitive processing (e.g. working memory, verbal intelligence) are enhanced and usually there is a correlation between the years of musical training and the enhancements. Additionally, there seems to be a neuroplastic change or modification created by the amount of musical exposure (Koelsch, 2005; Kraus & Chandrasekaran, 2010; Moreno & Bidelman, 2014; Patel, 2011; Seither-Preisler et al., 2014; Serrallach et al., 2016).

In their study, Seither-Preisler et al. (2014) and Serrallach et al. (2016) tested (1) in how far playing an instrument enhances the neural efficiency of auditory information encoding in the developing brain and (2) if children with musical training have enlarged HGs and faster auditory-evoked responses than children without training. They used both MEG and MRI scans to interpret auditory processing differences between musically-trained children and those without. The authors discovered that children who received musical training and had a higher index of musical practice (IMP) possessed enlarged right HG and, most interestingly, the volumetric measurements of the HG remained stable over the period of additional training. The authors therefore concluded that individuals are born with some kind of genetic predisposition for music, i.e., an enlarged HG and better auditory processing, which subsequently leads to intrinsic motivation and thus to the development of the exceptional ability. Sixty-percent of this ability was calculated to be due to a genetic predisposition, 40% due to environmental influences, supporting the innate aspect of aptitude.

Directly comparing language learning ability and musical ability, Milovanov (2009), Milovanov & Tervaniemi (2011) and Milovanov, Huotilainen, Velimäki, Esquef, & Tervaniemi (2008), Milovanov et al., 2009, Milovanov, Pietilä, Tervaniemi, & Esquef (2010), investigated Finnish native speakers and found significant relationships between musical aptitude and better second language pronunciation skills. They thus concluded that "regular music practice may also have a modulatory effect on the brain's linguistic organization" (Milovanov & Tervaniemi, 2011, p. 1). These results were also confirmed for Spanish and Japanese learners. Vangehuchten, Verhoeven and Thys, (2015) found a significant relationship between English pronunciation skills in Spanish native speakers and their musical skills, and Dolman and Spring, (2014) focused on Japanese learners and found that excellent skills in specific musical abilities, such as pitch, loudness and rhythm, correlate with better pronunciation in English (distinction and production between problematic sounds). Another study by Slevc and Miyake (2006) found a consistent relationship at least between musical aptitude and phonological aspects of linguistic ability, however not between other linguistic abilities such as syntactic or semantic skills.

Investigating language aptitude in more detail, Golestani, Molko, Dehaene, LeBihan, and Pallier (2007) reported correlations between an abnormal asymmetry of the planum temporale (the structure leading from the auditory regions towards the parietal lobe) and poor verbal skills. Wong, Perrachione, and Parish (2007) and Reiterer, Berger, Hemmelmann, and Rappelsberger, (2005), on the other hand, found more activation in less skilled learners in contrast to the better learners.[4] Focusing on a similar area as will be dealt with in the second part of the chapter, Golestani, Price and Scott (2011) undertook a project in which they examined the brain structure of expert phoneticians in order to find out whether any differences in language-relevant regions in the brain could be found between the experts and a control group. Interestingly, the size of the left inferior frontal gyrus (pars opercularis) correlated with the years of experience. Additionally, they found that the expert phoneticians had multiple or split transverse gyri in their left-hemispheric auditory cortex (which corresponds to HG) in comparison to the control group. In a vast project, Dogil & Reiterer (2009) tested German native speakers in a variety of areas to find possible correlations between foreign language pronunciation aptitude, working memory, language proficiency and language aptitude assessment. Phonological working memory (especially digit span forward) strongly correlated with L2 proficiency in English and neuroimaging methods revealed that talented, highly proficient learners develop more efficient processing networks for language. Individuals who had a very high ability in Hindi speech imitation also had excellent results in working memory tasks, higher scores in the MLAT test and also very good results in the English imitation tasks. Individuals with high pronunciation aptitude in English also had higher MLAT scores, liked acting, and had a higher openness to

[4] Such findings are in accordance with research on intelligence and the neural efficiency hypothesis stating that experts in a certain field have a better pruned network activating less regions necessary for accessing the skills (Neubauer & Fink, 2009).

experience, as well as more time spent abroad and more years learning English. Furthermore, musicality correlated positively and significantly with the pronunciation talent score and MLAT V.

To sum up, correlations between the variety of skills investigated have already been confirmed in recent studies but very few studies have actually included neuroanatomical measurements. The following research study aims at providing a clearer picture of the possible neuroanatomical correlates of musicality, working memory capacity and language learning ability in German speakers.

2 The Neuroanatomical Correlates of Foreign Language Aptitude: A Research Study

The main aim of this project was to investigate the auditory cortices of the individuals who took part in Dogil and Reiterer's study in 2009. We wanted to see whether we could confirm the correlations found between the most gifted and least gifted individuals (N = 30; 13 m/17f) and more importantly, we wanted to have a closer look at the primary auditory cortices of these individuals. Our main aim was thus to find the neuroanatomical correlates of musicality and language aptitude in a predefined group of German-speaking adults.

2.1 Methodology

The variables assessed were working memory (digit span, backward span, nonword span), musicality (AMMA), language aptitude (MLAT), speech imitation ability (Hindi) and other extra-linguistic factors (singing, number of instruments and languages learnt). The semi-automatic morphometric analysis developed by Peter Schneider (2002), Seither-Preisler et al. (2014), Serrallach et al. (2016), and Benner et al. (2017) was used to analyze the primary auditory cortices of the aforementioned subjects with the aim of revealing interesting insights into the brain of linguistically gifted individuals with varying degrees of musical ability.

2.1.1 Participants

All participants (*N* = 30; 13 male/17 female) were monolingual German native speakers between 20 and 40 years of age (*M* = 26.77, *SD* = 4.95). They had all commenced acquiring their first foreign language, namely English, at 10 ± 1 years of age. All participants were right-handed and were all either bachelor/master students or had achieved positions at an institution of higher education in different parts of Germany. None of the participants showed any medical condition or neurological disorder and they participated voluntarily in the research study.

Table 1 The parts of the MLAT being used as measures for language aptitude (Carroll, 1958 1962; Carroll & Sapon, 1957)

	Name	What is being tested?
Part 3	Spelling clues	Sound-symbol association ability and vocabulary knowledge – correct synonyms of disguised words have to be selected (multiple choice)
Part 4	Words in sentences	Grammatical sensitivity – components of sentences have to be identified (grammatical function) and related to elements in other words
Part 5	Paired associates	Associative rote memory – as many words in Kurdish have to be memorized as possible (presented with English translations)

2.1.2 Language Aptitude Testing

The participants were first classified as high or low aptitude speakers according to two scores, namely an English pronunciation score (based on the reading of the North Wind and the Sun) and a Hindi language aptitude score (speech imitation skills). To receive these scores, native speakers (of English and Hindi) rated the performance of the participants. In the Hindi test, participants had to repeat words and sentences in Hindi without prior knowledge of Hindi. Subjects outside one standard deviation from the mean (score ranging from 0 to 10) were classified as high and low aptitude speakers. An overall talent score was further calculated from the performance on both tests. Additionally, parts of the MLAT, namely parts III, IV and V were administered (see Table 1). The main focus was put on speech imitation aptitude and individuals were categorized into high and low talent groups according to this score.

2.1.3 Musicality Assessment

Musicality tests applied were AMMA tonal (pitch discrimination) and rhythmic (rhythm discrimination) part (Gordon, 1980, 2001) to assess the subjects' potential in the musical domain. Moreover, a questionnaire was given to subjects for assessing the number and types of instruments played and their liking and ability of singing.

2.1.4 Working Memory Skills

To assess subjects' working memory capacity, three different tests measuring complex and simple working memory skills were used, namely digit span forward, digit span backward and non-word span. All participants were given two chances for the same number of digits/non-words, i.e., if the first attempt of repeating a certain amount of digits failed, the subjects heard another set of the same amount of digits to repeat. Only if both attempts were incorrect, the test was stopped at that point and no points were given. Per correct series, the subjects received one single point.

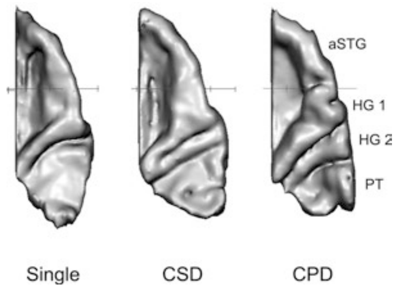


Fig. 2 3D reconstructions of the three types of HG distinguished in the analysis. Examples are given (left to right) for (1) single gyrus (SG), (2) common stem duplication (CSD), and (3) complete posterior duplication (CPD) (*aSTG* anterior superior temporal gyrus, *PT* planum temporale)

2.1.5 Morphometric Analysis

For the neuroanatomical analysis, the MRI scans of the 14 highest-scoring participants and the 16 lowest-scoring participants (based on the Hindi score) were analysed. T1-weighted structural magnetic MRI (Siemens, Magnetom SonataVision, 1,5 Tesla, software version: syngo MR 2004A, 176 DICOM slices, sagittal orientation, slice thickness 1 mm) had been performed for the study in 2009 and were now analysed to investigate the anatomy of the auditory cortex. Three-dimensional grey matter surface reconstructions of the primary auditory cortex (Heschl's gyrus) and the planum temporale were analysed using a standardized individual approach. This allows for a closer look at the morphometry and gyrification patterns in the segmented regions (Schneider et al., 2002, 2009; Serrallach et al., 2016). Brain Voyager software QX 2.8 (Brain Innovation, B.V, Maastricht, NL) was used for the segmentation of the aforementioned auditory-related areas. Pre-processing steps included the adjustment of brain images in contrast and brightness, as well as a correction for inhomogeneity and a rotation in direction of the antero-posterior commissural line. Normalization in stereotactic space (Talairach & Tournoux, 1988) had to be carried out to be able to compare the results. In the process of segmentation, the superior temporal plane, including HG, the anterior superior temporal gyrus (aSTG) and the planum temporale, was segmented into sagittal MRI slices along the lateral fissure using the standard definition of the landmarks of AC. A comparison of the three types of HG distinguished in this study are given in Fig. 2.

3 Results

3.1 Behavioral Results

The statistical analysis of the behavioral testing results was conducted using IBM SPSS 22. The descriptive results of the tests applied in the study shall be presented briefly.

Subjects received between 2.72 and 7.74 points (min. achievable: 0, max. achievable: 10) in the Hindi speech imitation task ($M = 4.81$, $SD = 1.64$) assessed by a large number of Hindi native speakers. The number of instruments subjects had learnt ranged from zero to three ($M = 1.23$, $SD = 0.97$), but most participants played one instrument. The number of foreign languages acquired differed substantially, ranging from one to nine ($M = 2.59$, $SD = 1.72$). The English pronunciation score ranged from 0 to 10 although none of the participants were rated below a score of 2 ($M = 6.40$; $SD = 1.72$). AMMA tonal results ($M = 28.72$, $SD = 5.68$) differed slightly from AMMA rhythm results ($M = 31.10$, $SD = 4.61$) and the total AMMA score, subsuming both subtests, ranged from 42 to 79 ($M = 59.8$, $SD = 10.05$). The measures for working memory capacity, digit span forward ($M = 9.59$, $SD = 1.88$) and digit span backward ($M = 8.76$, $SD = 2.13$) gave similar results, though subjects had better performances in the forward task. The digit span backward scores ranged from 4 to 13, overall slightly higher than the span for the non-word task ($M = 7.55$, $SD = 1.74$), in which subjects scored between 5 and 11 points. Considerable variability was found in the MLAT total raw with a range from 49 to 109 points ($M = 83.41$, $SD = 14.23$), consisting of scores for part III ($M = 36.69$, $SD = 8.62$), part IV ($M = 29.28$, $SD = 5.58$) and part V ($M = 17.31$, $SD = 5.09$).

The number of instruments played by a subject correlated positively and significantly with a large number of other variables, namely the total AMMA musicality score ($r = .432$, $p = .019$), the Hindi speech imitation score ($r = .385$, $p = .035$), the total talent score ($r = .454$, $p = .012$), the non-word repetition score ($r = .369$, $p = .049$), the overall MLAT score (raw) ($r = .379$, $p = .043$) and most strongly the MLAT IV (measuring grammatical sensitivity) ($r = .493$, $p = .007$). In contrast, no correlations could be found for the number of foreign languages spoken by a subject and any behavioral test scores. In other words, there was no difference in performance according to the number of languages spoken by the participants. The speech imitation score, i.e., the Hindi score, correlated positively and significantly with the English pronunciation score ($r = .395$, $p = .031$), digit span forward ($r = .432$, $p = .019$), non-word repetition ($r = .371$, $p = .047$), the MLAT total score ($r = .452$, $p = .014$) and also one component of the MLAT total score, namely the MLAT IV ($r = .399$, $p = .032$).

The English pronunciation score did not only correlate positively and significantly with the Hindi speech imitation score, but also showed significant correlation with the MLAT total raw score ($r = .677$, $p = .00006$) and two components of the MLAT, namely part IV ($r = .486$, $p = .007$) and III ($r = .673$, $p < 0.001$). The overall talent score subsumes both the Hindi and the English score and thus positive, significant correlations could be found with the AMMA total score ($r = .368$, $p = .05$), even more so the AMMA rhythm score ($r = .383$, $p = .04$), non-word repetition ($r = .384$, $p = .04$) and all subparts of the MLAT and the total MLAT score ($r = .713$, $p < 0.001$). As expected, the tonal and rhythmic parts of Gordon's AMMA test correlated positively with each other ($r = .938$, $p < 0.001$) and the number of instruments played.

A t-test based on the distinction between talented and non-talented subjects as ranked by the Hindi speech imitation score was conducted. Significant differences between talented and non-talented individuals were found for the number of played instruments ($t_{(28)}$ = −2.32, p = .028), English pronunciation score ($t_{(28)}$ = −2.1, p = .045), digit span forward ($t_{(27)}$ = −2.73, p = .011), non-word repetition ($t_{(27)}$ = −2.5, p = .017) and MLAT total raw scores ($t(27)$ = −2.27, p = .032).

3.2 Results of the Neuroanatomic Analysis

The neuroanatomic results of the manual segmentation of the auditory cortices revealed interesting insights into the neurological underpinnings of language aptitude and musical skills. The total number and percentage of types of HG found in the participants is provided in Table 2. The results clearly indicate that individuals with high speech imitation aptitude (as explained in the previous paragraphs) have more complete posterior duplications of their Heschl's gyri in both hemispheres but in particular in the right hemisphere. Those who had excellent scores in the Hindi aptitude testing do not only have one single Heschl's gyrus, like the subjects of the second group, but they have two equally prominent Heschl's gyri. Even if the trend of more complete (i.e., going from the lateral to the ventral end) posterior duplications is exceptionally stable in the right hemisphere, also the left hemisphere showed more frequently complete posterior duplications in highly gifted individuals in contrast to the comparison group – the number was generally just too limited to draw conclusions (only 4 of the gifted 14 had a CPD). The neuroanatomical results clearly indicate that individuals with high speech imitation aptitude in the Hindi testing, and also individuals with very high scores in the AMMA testing, showed more complete posterior duplications of their HG in the right hemisphere. Figure 4 displays three exemplary individuals of each group (very high and very low scores in the Hindi speech imitation task) (Fig. 3).

In order to verify the significance of the observations that had been made beforehand, a one-way ANOVA for subjects displaying one of the three following morphological HG characteristics in their right hemisphere, namely (1) single gyrus (SG), (2) common stem duplication (CSD) and (3) complete posterior duplication (double gyrus; CPD), was performed. A significant group difference could be observed for the Hindi speech imitation score ($F(2,27)$ = 5.9, p < .01, part. η^2 = .30. Subjects with a CPD achieved significantly higher scores (6.1 ± 1.2) than subjects with a SG (4.1 ± 1.4; p = .009) and subjects with a CSD (4.4 ± 1.7; p = .034) (see Fig. 4).

Table 2 Frequency of types of HG in right and left hemispheres in subjects

		RH	LH
	Total number (%)	30 (100%)	30 (100%)
Types of HG	Single	12 (40%)	23 (76,7%)
	CSD	9 (30%)	3 (10%)
	CPD	9 (30%)	4 (13,3%)

RH right hemisphere, *LH* left hemisphere, *CSD* common stem duplication, *CPD* complete posterior duplication

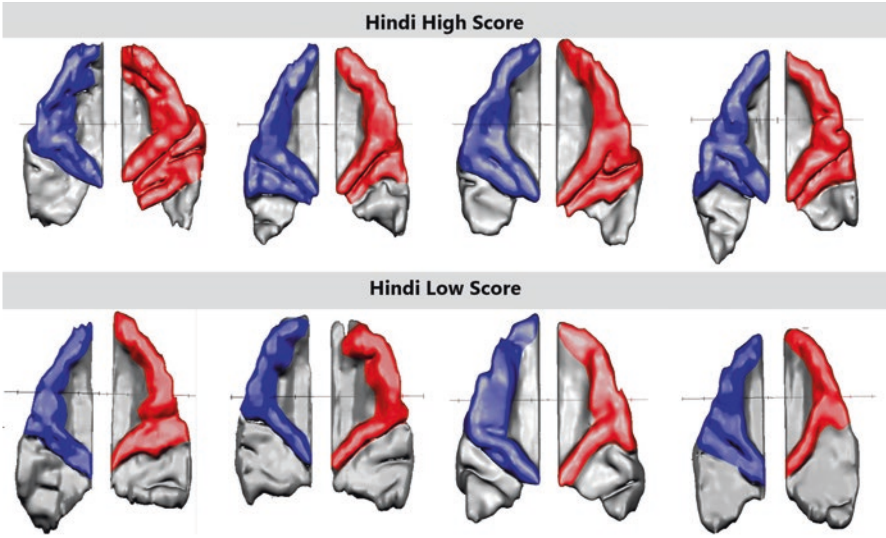
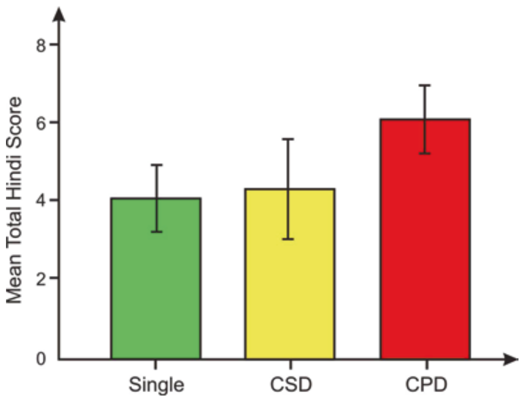


Fig. 3 A comparison of the auditory cortices of individuals with very high (above) and very low speech imitation scores (below). (Red – right hemisphere, Blue – left hemisphere)

Fig. 4 Results of the one-way ANOVA comparing mean total Hindi score (overall range: 0–10) with the three types of HG in the right hemisphere. Individuals with CPD scored significantly higher in the Hindi testing in comparison to subjects with SG or CSD in the right hemisphere



Similar results were found for the AMMA test. The mean of the total AMMA score in the right hemisphere for SG was 56.1 ± 10.2, for CSD 55.5 ± 7.6, and for CPD 68.9 ± 5.7. Individuals with CPD achieved significantly higher scores than subjects with SG and CSD ($F(2,26) = 7.95$, $p < .01$, part. $\eta^2 = .38$). Individuals with CPD in the right hemisphere achieved significantly higher scores than subjects with SG and CSD. The mean of the total AMMA score for SG was 56.09, $SD = 10.21$, for CSD 55.53, $SD = 7.55$, for CPD 68.89, $SD = 5.67$. The results of the statistical analysis $F(2,26) = 7.95$, $p < .01$, $\omega = .57$, confirm that subjects who possess CPD in the right hemisphere performed significantly better in the AMMA musicality test and seem to have higher musical ability. No significant difference could be found for SG and CSD, though.

4 Discussion

The primary aim of this study was to investigate the following research questions:

1. To what extent does language aptitude depend on other cognitive variables such as musicality and working memory?
2. Can the structure of the primary auditory cortex (more specifically Heschl's gyrus) be seen as a neuroanatomical marker of language aptitude?

These two questions will be addressed and put into the context of previous research in the following subsections.

4.1 Language Aptitude

The results of the t-tests revealed that the more gifted participants according to the Hindi speech imitation score played more instruments, demonstrated higher proficiency in English pronunciation, scored higher in the working memory tasks and also had better results in the total MLAT score. Furthermore, they had more double gyri in the right hemisphere in contrast to the less gifted ones. Nevertheless, the concept of language aptitude is very complex (see introductory chapter of this book) and it is of utmost importance to discuss the relevance of the findings for contemporary language aptitude research and future studies.

There is no doubt that the two main tests applied in this research project, namely the Hindi speech imitation task and the English pronunciation proficiency test, measure very different components of foreign language learning ability. The Hindi test is a speech imitation score, which basically requires the imitation of foreign speech material (Dogil & Reiterer, 2009; Jilka, 2009) and thus assesses basic perception and production skills of an unknown language. The English pronunciation testing, on the other hand, gives an overview of pronunciation proficiency in an already acquired language. The difference between the two is clearly that while the first measures phonetic/phonemic coding ability (a major component of foreign language aptitude), the second could be influenced by a large amount of factors and marginally depends on language aptitude. Owing to these considerable differences, it is essential to differentiate between aptitude and proficiency here (and not argue that both indicate high language aptitude) but still, a moderate, positive correlation was found between the two scores ($r = .395$, $p = .031$). On the one hand, the Hindi score can be seen as a predictor for the English score given that phonetic coding ability is a core component of language aptitude and should hence result in an excellent pronunciation in any language being acquired. One major concern with English in this case, though, might be that several subjects had spent time in an English-speaking country or had even pursued a degree in English. Furthermore, English is introduced as the first foreign language for almost all children in Germany, which means that acquiring a native-like pronunciation is already supported from childhood on. This could be an explanation for the fact that a number of individuals with lower scores in the Hindi test scored much better in the English pronunciation test (a possible explanation for the moderate correlation). As will be explained in more detail, we argue that language aptitude is defined as a rather innate capacity that develops over time but remains quite stable. We therefore argue that it is the Hindi scores in this case that predict the English scores, i.e., the better individuals are at decoding, retaining and reproducing unknown speech material, the easier it is for them to develop excellent pronunciation skills in a given language. Still, given an outstanding environment, stays abroad and numerous years of practice, high proficiency may be achieved despite moderate phonetic coding ability.

With respect to the results of the MLAT, clear significant positive correlations were found between the English pronunciation, results of the MLAT (particularly parts IV and III) and also between the Hindi speech imitation and the MLAT overall score. An interpretation of these scores is twofold: first we could argue that both aforementioned tests (Hindi and English) measure, even if only to a certain and especially differing extent, language learning ability, i.e., language aptitude. Assuming that the subtests of the MLAT (phonetic ability in part III and grammatical sensitivity in part IV) are excellent indicators of these specific components of aptitude, we would expect a rather high, positive correlation between the Hindi speech imitation and MLAT part III since both are considered measures of phonetic coding ability. This was not the case, though. An often-cited and quite obvious disadvantage of the MLAT is the fact that the test has to be taken in English. This certainly gives individuals with higher proficiency in English a clear advantage and they could achieve high scores only because of their English skills and not their language aptitude. Furthermore, it gives individuals with native languages similar to English (the test stimuli involve English structures) another obvious advantage. This could be the case with our German speakers and this could be a serious issue. This obstacle could also be overcome by using more recently developed language-independent aptitude tests, such as the LLAMA language aptitude test (Meara, 2005). In fact, the LLAMA test has gained considerable popularity in the past years and it has thus been more and more frequently applied (Artieda & Muñoz, 2016; Granena & Long, 2013; Kepinska, de Rover, Caspers, & Schiller 2017a, 2017b; Kepinska, 2017; Rogers et al., 2016) and also many studies of this book applied the LLAMA. To sum up the results regarding the MLAT, it could be merely a good proficiency in English that is responsible for outstanding results in parts of the MLAT, at least this is very likely in part III and we therefore cannot really argue that the MLAT captures language aptitude to a full extent in this study.

However, the results of the MLAT demonstrate that the language proficiency testing (English) also correlates with grammatical sensitivity and we assume that this should be rather a result of aptitude than of proficiency. Language analytic ability, an umbrella term including abilities in grammatical sensitivity, is an important component of language aptitude recently investigated by Kepinska (2017) and Kepinska et al. (2017a, 2017b). We suggest that grammatical sensitivity is of high significance for the learning of foreign languages and will need to be dealt with more explicitly in future studies. In our case, though, we focused on foreign language pronunciation ability (Jilka, 2009) and research has shown that both do not always necessarily go hand in hand.

As the vocabulary learning task of the MLAT (part V – associative memory) did not correlate with the English pronunciation score, we propose first that this is partly due to the fact that excellent English skills were not that important in this test. Vocabulary learning tasks further require excellent memory skills and we therefore assumed that the Hindi performance somehow goes hand in hand with what is measured by MLAT V. Verbal working memory skills (attributed to the phonological loop) are essential for novel vocabulary learning (Atkins & Baddeley, 1998; Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1990; Gathercole, Service, Hitch, Adams, & Martin, 1999) but MLAT V could surprisingly not be linked to any working memory task. With respect to the overall MLAT score, the Hindi test correlated positively with the MLAT overall score. This could be interpreted as a result that the Hindi speech imitation measures language aptitude in broader terms and the statistical significance between the Hindi test and MLAT part III and IV were compensated for by the rather strong relationship with part V. As a consequence, only a relationship between the overall raw score and the Hindi score could be found. A larger sample would definitely be needed in order to be able to confirm these hypotheses.

One result that needs separate discussion is the fact that the number of languages spoken by the participants (i.e.., the languages acquired beforehand) could not be linked to any other score. It has often been argued that the more languages one has already learned and master, the better one is at learning new languages, or the better English pronunciation skills and speech imitation ability should be. Vice versa, individuals with very high language learning ability could be more likely to learn more languages as they acquire foreign languages much more easily. It came as a surprise that this was not the case but various reasons could be linked to this lack of correlation. First of all, not everybody, regardless of their aptitude for acquiring a foreign language, wants to learn or has the chance to learn numerous languages (possibly due to a lack of time, opportunity or necessity). Secondly, having learnt a considerable number of foreign languages does not necessarily imply that it was easy for a person to learn this language – it would be a mistake to draw this conclusion without having asked the participants. Also, the sample was limited and unfortunately, it could not be controlled in any form how well the participants had learnt the languages. Their proficiency was never assessed as part of this study and often self-assessed proficiency is misleading since individuals hardly distinguish between different skills (e.g., writing, speaking, reading) and their view of themselves may be completely wrong. Even if our subjects had stated that their level in French was A2 this would not have provided us with sufficient information to include their proficiency as a variable. It would have been necessary to gather information on proficiency (grammar, vocabulary, pronunciation) in all foreign languages and concretely define what is meant by proficiency in our study.

To get back to the point, the inability to link the number of languages having been acquired to language learning ability strongly supports our claim that language aptitude is a rather innate and inflexible capacity. The initial claim described language aptitude as a trait that cannot be altered through learning or practice (Carroll, 1990; Stansfield & Reed, 2004). Despite the impact foreign language learning has been suggested to have on cognition and a variety of other skills, there seems to be no evidence from our study supporting the claim that previous language experience impacts foreign language aptitude (Thompson, 2013). As we suggest that language aptitude is a potential gift before the acquisition of any language, speaking two or nine foreign languages should not have any influence. Furthermore, even though the stability and fixed nature of the concept of language aptitude have been questioned in past decades (Klein, 1995; Sáfár & Kormos, 2008; Thompson, 2013), aptitude may not be such a highly dynamic construct as most recently proposed.

Finally, the number of instruments played by a subject and two tests of working memory capacity, namely digit span forward and non-word repetition, correlated positively and significantly with the Hindi speech imitation score. Since it is not a secret that musical ability may deeply impact foreign language learning, we see the positive relationship between speech imitation and the number of instruments played by a subject as a strong support for this claim. We agree that playing an instrument certainly enhances auditory processing in an individual. For this reason, we also expected a strong relationship between the other musicality scores (AMMA) and the Hindi speech imitation test (for details, see next paragraphs). Also the very positive results linking language learning ability to working memory capacity are in accordance with what recent research has confirmed. The details on working memory capacity and musicality are discussed in the following sections.

4.2 Musicality

As discussed quite extensively in the literature, the number of instruments played by an individual has often been assumed to have a major influence on a variety of related skills, such as foreign language learning (Milovanov et al., 2008; Nardo & Reiterer, 2009). This process is termed positive transfer and Kraus and Chandrasekaran (2010) have therefore made a good point describing music as a resource that results in auditory fitness. Our results definitely support the known relationship between the musical domain and linguistic skills. Generally, those subjects who achieved high results in the overall talent score (Hindi and English combined in one score), played more instruments and performed significantly better in the musicality tests. This finding is in accordance with very recent research studies exploring the two domains (Christiner & Reiterer, 2013; Dogil & Reiterer, 2009; Fonseca-Mora, Toscano-Fuentes, & Wermke, 2011; Milovanov et al., 2008, 2009, 2010; Seither-Preisler et al., 2014). It was quite surprising and somewhat unexpected though that only the total talent score showed a highly significant relationship with the AMMA test, but not the two core tests respectively. Performances combined in English and Hindi correlated positively and significantly with AMMA rhythm and the overall AMMA score. The Hindi score alone only showed a positive trend when we looked at its relationship with the AMMA total score, which suggests that the correlations should be statistically significant in a larger sample. We assume that musical ability enhances the processing of language and despite the fact that there was only a positive trend, we are sure that phonetic coding ability and musicality profoundly influence one another.

Furthermore, the number of instruments played by the participants also yielded a strong, positive correlation with the Hindi score, which provides further evidence for the relationship between music and language. As expected, the two subparts of the AMMA correlated significantly with each other, suggesting that musical ability brings along very good ability in different musically-related tasks, in this case rhythm and pitch discrimination. The two AMMA tests also correlated positively and significantly with the number of instruments played and this clearly indicates that individuals who learn to play more instruments have an advantage in auditory discrimination, i.e., a well-trained auditory cortex (be it present before learning to play the instruments or not) (Kraus & Chandrasekaran, 2010). Conversely, this can be seen as another confirmation that the AMMA test is an excellent measure of musical ability and that high musical ability leads to learning of instruments and vice versa. The fact that only a moderate correlation between the two could be found can be explained by the simple fact that the amount of time and practice subjects had to play an instrument had not been taken into account and assessed properly. Future research will need to take these factors into consideration to specify the influence musical practice and intrinsic motivation might have.

To sum up, future studies will need to spend more time investigating the concept of musical aptitude or musicality and use a larger variety of measures to fully grasp the construct. It seems true that more factors and variables need to be taken into account and calculating a musicality index would therefore be a good option (see Seither-Preisler et al., 2014; Serrallach et al., 2016). Even though the AMMA is usually taken as a measure of musical ability generally, there is no doubt that a musicality index and different measures on musical ability are needed to shed more light onto this issue.

4.3 Working Memory Capacity

The claim that working memory hugely impacts foreign language aptitude is definitely not far-fetched (Miyake & Friedman, 1998; Wen, 2012; Wen & Skehan, 2011; Wen et al., 2017). By challenging this assumption, studies have found that speech imitation skills rely heavily on working memory (Biedroń, 2012; Ellis & Sinclair, 1996; Kormos & Sáfár, 2008; Linck et al., 2013; Miyake & Friedman, 1998; Sáfár & Kormos, 2008) and quite in accordance with these results, we found a positive relationship between speech imitation skills and the different tasks assessing working memory skills. The three tasks we used were digit span forward, digit span backward and non-word span and all three of them correlated positively and significantly with the Hindi score. Not surprisingly, the non-word span showed by far the highest correlation in this regard. It is interesting to note, though, that the three working memory scores did not show correlations with any other variable, e.g. AMMA, number of instruments, number of languages or MLAT. Only the Hindi score showed strong correlations.

We first hypothesized that phonetic coding ability should be measured most successfully by using the Hindi speech imitation task. Interestingly, studies have suggested that the non-word span can be used for assessment of linguistic difficulties, such as for specific language impairment (Botting & Conti-Ramsden, 2001; Coady & Evans, 2008), which supports our hypothesis that it might serve well for differentiation between particularly high or low language learning ability (assuming specific language impairment portrays the other end of the continuum). Both the Hindi speech imitation task and the non-word span require the decoding of unfamiliar speech stimuli, their storage in working memory and the ability to correctly reproduce them. Slight differences between the two are that the non-word span gets more difficult with each level but uses the same stimuli, i.e., while the working memory load increases, the difficulty rather stays the same. This is not the case with the Hindi task since the stimuli always change but do only slightly get more difficult or complex (changing from long words to short sentences only adds marginal difficulty for the participants). Another difference concerns the nature of the stimuli. The Hindi speech input is that of a natural language and does not resemble a particular type of language (the non-word span syllables are German non-words). Although there are slight differences, they both use speech material basically consisting of simple syllables and require the same processing steps. In sum, we argue that both heavily rely on working memory capacity and we propose that both tests are equally useful measures of working memory capacity on the one hand, and phonetic coding ability on the other. We therefore agree with other researchers that non-words tests should be included in language aptitude testing batteries since they are good additional measures of phonetic coding ability (Chan, Skehan, & Gong, 2011).

As already mentioned, digit span forward and backward both yielded strong correlations with the Hindi score, not with the English pronunciation score though. This is a little surprising since other studies (Biedroń, 2012; Kormos & Sáfár, 2008; Van den Noort et al., 2006) have shown that high language learning ability and successful foreign language acquisition (proficiency) correlate with working memory tasks of differing complexity. We do not have a concrete explanation for this but since we argue that the English pronunciation score is much more a measure of proficiency than ability, the claim that language aptitude depends on working memory still holds true.

In summary, we conclude that both simple and complex working memory skills are required to imitate foreign speech material, i.e., for the skill termed phonetic coding ability. Our results thus support the hypothesis that working memory is a core component of foreign language aptitude (Wen, 2016; Wen et al., 2017) but we do not completely agree with the hypothesis that working memory may be seen as an equivalent to foreign language aptitude as recently suggested (Linck et al., 2013; Wen, 2012, 2016).

4.4 Neuroanatomic Markers for Language Aptitude and Musicality

Analyzing brain structures to explore the neural underpinnings of certain abilities, such as foreign language learning ability, has definitely gained increasing popularity in the past years. Few studies have attempted to link concrete regions to concrete skills anatomically with all the focus being put on the functional properties of language learning. According to Berken et al. (2015), structural variation in the brains of individuals may indeed provide a partial explanation for greater language aptitude. Learning novel elements of a language, such as learning tonal pitch contrasts and phonetic differences (Golestani & Zatorre, 2009), as well as perceiving and producing novel speech sounds (Golestani & Pallier, 2006; Golestani, Paus, & Zatorre, 2002), can provide us with interesting information as to which regions are important for these processes (Berken et al., 2015). Language has predominantly been ascribed to the left hemisphere and also findings regarding the language-related functions of HG have emphasized the role of HG on the left side (Golestani & Pallier, 2006; Golestani et al., 2007).

The neuroanatomic analysis performed by the authors clearly showed that those individuals who had higher Hindi scores also had more CPDs of HG in the right hemisphere. Of course, this is somehow contradictive with regard to the importance the left HG has received for its role in linguistic processing and thus foreign language learning. Nevertheless, the statistical results clearly indicated this positive relationship between more gyri on the right side and better speech imitation skills. What is quite striking, though, is that the AMMA score also showed a particularly strong relationship with HG duplications in the right hemisphere. This result is particularly remarkable and challenging to explain since the AMMA score and the Hindi score did not correlate as expected. In other words, the duplications cannot be attributed to the two scores in combination but seem to be important for each respectively. Two topics will be specifically addressed to discuss these findings accordingly. First of all, could it be that the results of the neuroanatomic analysis suggest more than just a positive relationship between language learning ability and music? It seems vital to specify the nature of this relationship, the role of the auditory cortex on the two skills and even more importantly, the impact this may have on HG functioning in the left and right hemisphere. Second, up to date we can only hypothesize about the concrete function of CPDs in HGs and very little research, almost none in the linguistic domain, have addressed this issue. Leaving aside the hemispherical differences, we can only make guesses as to why some individuals have more than one gyrus and in how far this influences auditory functioning and thus the acquisition of linguistic and musical skills.

Our primary aim when conducting this analysis was to find neuroanatomical markers of language aptitude. Doubtlessly, the auditory cortex is only a starting point and there are certainly dozens of regions that may influence foreign language aptitude (e.g., supramarginal gyrus, angular gyrus, inferior frontal gyrus). As already mentioned, language has been claimed to be predominantly left-lateralized and we therefore expected to find differences mainly in the left hemisphere as in other studies (Dogil & Reiterer, 2009; Golestani & Pallier, 2006; Golestani & Zatorre, 2009; Golestani et al., 2011, 2002, 2007; Hu et al., 2013; Reiterer et al., 2011; Warrier et al., 2012; Wong et al., 2007). Contrary to our expectations, the variability we found in the left hemisphere was so marginal that no statistical analysis could be performed (see Table 2) and a much larger sample would have been needed to find interpretable variability in the left hemisphere. The results of the right hemisphere are quite remarkable though and shall receive sufficient attention in the following paragraphs.

It is known that musical ability heavily relies on the right hemisphere and also recent research has proved that HG on the right side is essential for musical processing. The results clearly indicate that the shape of the auditory cortex and the fact that an individual has more than one HG in the right hemisphere are linked to both musical and phonetic coding ability. Only the individuals with a CPD (i.e., two complete gyri) in the right hemisphere had significantly higher scores in the AMMA test and in the Hindi task. The individuals with SG and CSD (not counted as two complete gyri) had substantially low results in both. Why did we only find differences in the right hemisphere, though? Is it necessary to have two HGs in the right hemisphere to have a considerable advantage in both musicality and phonetic coding? And if we assume that individuals with a double HG have a better developed auditory cortex, why is it that both language and music seem to be so heavily influenced by it? We thought of a number of possible explanations and we would like to share a couple of them.

We are quite certain that we would have discovered a stronger relationship between language aptitude and musicality in a larger sample. Our Hindi task requires excellent use of the articulators to reproduce foreign speech and a considerably well-developed auditory cortex to hear the subtle differences in the speech input provided. Could it be that only phonetic coding ability, i.e., this one component of foreign language aptitude, is highly dependent on (1) musical skills or (2) auditory processing in the right hemisphere? If we argue that phonetic coding ability depends on auditory processing of music-relevant features known to be processed in the right hemisphere, this could provide us with an explanation why only the right hemisphere showed structural variation in the subjects. A question that immediately arises with this regard is that of whether differences in the auditory cortex can be attributed to language aptitude or musical ability. What if it were a combination of both or of more general ability that ultimately influences both? It could be that we just found a confirmation of the importance of CPDs in the right hemisphere for processing of musical features and since Hindi speech imitation requires non-speech processing expertise, we found a similar relationship between the two. This hypothesis sees musical processing as the predominant skill and speech imitation as a result of that. Another explanation would be the inverse, meaning that the findings suggest a neuroanatomical marker for foreign language pronunciation aptitude and this innate marker also highly influences musical processing, ultimately leading to an outstanding ability in both domains. Adopting this view, high potential for learning foreign languages would be seen as the core skill which facilitates musical processing. Finally, the last option would be to speak of a general auditory processing potential that can influence both, depending on the stimulation and the preference of each individual. This could lead to high ability in one domain or even in high ability in both domains since we know that musicians are often also good language learners. These are only hypotheses and even if the one or the other seems more likely, it is also possible that the solution is a combination of all three approaches.

The differentiation between the functions of HG in the right and left hemisphere might lead to future discussions regarding the left-lateralized view on language processing. One major issue here is that we do not know whether auditory processing differences existed between the individuals with high and low aptitude. Even though language skills are more linked to the left hemisphere, it might be that the Hindi task specifically also addresses the right hemisphere (phonological decoding, not grammatical or semantically-related tasks) and this is the reason why we found differences in the right hemisphere only. Another explanation would be that auditory processing differs between highly gifted individuals and the low talent group, which we unfortunately cannot confirm in a neuroanatomic analysis. It would be necessary to apply fMRI or EEG/MEG to detect possible differences here, which could then consequently explain structural variation encountered. Quite recently an appealing study by Kepinska et al. (2017a, 2017b) and Kepinska (2017) on language analytic ability highlighted the significance of the right hemisphere for language aptitude. They focused on grammatical analytic ability and found that regions in the right hemisphere were at least equally important for that specific skills. The right hemisphere might thus be more important than initially assumed but more research will be needed to explore the involvement of the right hemisphere, in particular the right HG, in different aspects of language aptitude. Also, given the various regions in the brain that are essential for language processing, it will be essential to develop methods in order to structurally analyze other significant areas, such as the inferior parietal lobule or regions in the inferior frontal cortex that seem to be highly important (Dogil & Reiterer, 2009; Golestani et al., 2002, 2007, 2011; Hu et al., 2013; Reiterer et al., 2011).

Although considerable inter-individual differences have been found, individual brain morphology has been shown to be extremely stable (Seither-Preisler et al., 2014; Serrallach et al., 2016). This means, that structures such as gyri or sulci do not undergo change from childhood into adulthood and it has therefore been suggested that these differences are first and foremost not due to environmental influences or practicing behavior. Quite on the contrary, their stability seems to point towards a strong biological component, which may be genetic, prenatal, or very early post-natal (Schumann, 2004). Numerous studies have focused on neuroplasticity but none have shown so far that the main structure of the brain itself actually changes after the birth of an individual. Another difficult issue is that it is still unclear how the gross-morphological structural characteristics of the auditory cortex are related to functional activation patterns. In particular, the suspected advantage a CPD of HG has in an individual's brain and if and for what reason this affects language learning ability and musical ability remains to be uncovered in the future decades.

Surely, we at no point question that numerous areas in the brain are of greatest importance for language processing but there is no doubt that the auditory cortex is a major path of language to the brain. We are aware that the view of language aptitude has changed in the past decades and it is more and more frequently referred to as a dynamic construct that may indeed undergo change over time. Still, if we are able to determine neuroanatomical markers for foreign language aptitude and these markers have been found to be stable over time, this challenges the assumption that language aptitude is something that can be altered. If the structure of certain brain regions (we are not talking about plasticity here) has an influence on skills, we have to find out in how far predetermined structures play a role for the development of these skills. Surely, numerous variables influence the development of language aptitude but we support the claim that there is an extent of innate markers that remain to be found. Even though this might suggest that aptitude is definitely something an individual is born with and that cannot be learnt, we do not support the claim that language aptitude is necessary for acquiring foreign languages or achieving a high proficiency at all – this has never and will never be our aim. We are already working on similar investigations in children with differing degrees of musicality and language ability in order to confirm what has been found in this study. We also highly encourage other researchers to investigate language aptitude from an anatomical viewpoint and help uncover the neuroanatomic underpinnings of language aptitude.

5 Conclusion and Implications for Future Research

The results of our research lead us to the conclusion that neuroanatomical markers of phonetic coding ability, or more generally foreign language learning ability, can be linked to classical auditory areas known to be involved in music processing. This suggests that abilities in the domains of music and speech do not only overlap and correlate on a behavioral level, but might even share common grounds in the neuroanatomy of primary auditory areas – even if the concrete relationship cannot be specified yet.

The findings of this study question the directedness of the relation between music and language, the direct neuroanatomical overlap and the functional consequences caused thereby. Also, the role and importance of HG and the right auditory cortex, as well as structural variation more generally shall be highlighted within this chapter. The possible explanations for the results of our research project were dealt with extensively in the discussions section but the issue cannot be solved with the data acquired in this study.

Doubtlessly, larger groups of participants are needed to confirm these findings and more research needs to be done in particular with regard to structural variation in individuals. The hunt for the neuroanatomical markers of foreign language aptitude has only begun and interesting results await to be reported. Additionally, further exploring the relationship between music and language will surely have far-reaching consequences for education (school, university, foreign language learning) and our view of aptitude and giftedness.

References

Abdul-Kareem, I. A., & Sluming, V. (2008). Heschl gyrus and its included primary auditory cortex: Structural MRI studies in healthy and diseased subjects. *Journal of Magnetic Resonance Imaging, 28*(2), 287–299.

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481–509. https://doi.org/10.1017/S027226310808073X

Al-Shabatat, A. M. (2013). A review of the contemporary concepts of giftedness and talent. *International Interdisciplinary Journal of Education, 2*(12), 1336–1346. https://doi.org/10.12816/0002983

Artieda, G., & Muñoz, C. (2016). The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency. *Learning and Individual Differences, 50*, 42–48.

Atkins, P. W., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*(4), 537–552.

Baddeley, A. (2003a). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A. (2003b). Working memory and language: An overview. *Journal of Communication Disorders, 36*(3), 189–208. https://doi.org/10.1016/S0021-9924(03)00019-4

Baddeley, A. (2017). Modularity, working memory and language acquisition. *Second Language Research, 33*(33), 299–311. https://doi.org/10.1177/0267658317709852

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158–173. https://doi.org/10.1037//0033-295X.105.1.158

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47–90.

Baddeley, A. D., & Hitch, G. J. (2000). Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged? *Journal of Experimental Child Psychology, 77*(2), 128–137. https://doi.org/10.1006/jecp.2000.2592

Bear, M. F., Connors, B. W., & Paradiso, M. A. (Eds.). (2007). *Neuroscience. Exploring the brain*. Baltimore: Williams & Wilkins.

Benner, J., Reinhardt, J., Schneider, P., Stippich, C., Wengenroth, M., & Blatow, M. (2017). Prevalence and function of Heschl's gyrus morphotypes in musicians. *Brain Structure and Function, 222*(8), 3587–3360. https://doi.org/10.1007/s00429-017-1419-x

Berken, J. A., Gracco, V. L., Chen, J. K., Watkins, K. E., Baum, S., Callahan, M., & Klein, D. (2015). Neural activation in speech production and reading aloud in native and non-native languages. *NeuroImage, 112*, 208–217.

Besson, M., & Schön, D. (2001). Comparison between Language and Music. *Annals of the New York Academy of Sciences*, 232–258. https://doi.org/10.1111/j.1749-6632.2001.tb05736.x

Biedroń, A. (2011a). Personality factors as predictors of foreign language aptitude. *Studies in Second Language Learning and Teaching, 1*(1231), 467–489.

Biedroń, A. (2011b). Intelligence in gifted L2 learners. In M. Pawlak (Ed.), *Second language learning and teaching* (pp. 129–142). Berlin, Germany: Springer.

Biedroń, A. (2012). Memory abilities in gifted foreign language learners. In M. Pawlak (Ed.), *New perspectives on individual differences in language learning and teaching* (pp. 77–96). Berlin, Germany: Springer.

Biedroń, A. (2015). Neurology of foreign language aptitude. *Studies in Second Language Learning and Teaching, 5*(1), 13–40. https://doi.org/10.14746/ssllt.2015.5.1.2

Biedroń, A., & Pawlak, M. (2016a). New conceptualizations of linguistic giftedness. *Language Teaching, 49*(2), 151–185. https://doi.org/10.1017/S0261444815000439

Biedroń, A., & Pawlak, M. (2016b). The interface between research on individual difference variables and teaching practice: The case of cognitive factors and personality. *Studies in Second Language Learning and Teaching, 6*(3). https://doi.org/10.14746/ssllt.2016.6.3.3

Botting, N., & Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI). *International Journal of Language & Communication Disorders, 36*(4), 421–432.

Brown, H. D. (2006). *Principles of language learning and teaching*. White Plains, NY: Pearson Education.

Carroll, J. B. (1958). A factor analysis of two foreign language aptitude batteries. *Journal of General Psychology, 59*, 3–19.

Carroll, J. B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87–136). New York: Wiley.

Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *Linguistics, 11*(111), 5–14.

Carroll, J. B. (1990). Cognitive abilities in foreign language aptitude. In T. S. Parry & C. W. Stansfield (Eds.), *Language aptitude reconsidered* (pp. 11–30). Englewood Cliffs, NJ: Prentice Hall.

Carroll, J. B., & Sapon, S. (1957). *Modern language aptitude test*. Bethesda, MD: Second Language Testing.

Chan, E., Skehan, P., & Gong, G. (2011). Working memory, phonemic coding ability and foreign language aptitude: Potential for construction of specific language aptitude tests–the case of Cantonese. *Ilha do Desterro: A Journal of English Language, Literatures in English and Cultural Studies, 60*, 45–69.

Chi, J. G., Dooling, E. C., & Gilles, F. H. (1977). Gyral development of the human brain. *Annals of Neurology, 1*, 86–93. https://doi.org/10.1002/ana.410010109

Christiner, M., & Reiterer, S. M. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00874

Coady, J. A., & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 43*(1), 1–40.

Dogil, G., & Reiterer, S. M. (2009). *Talent and brain activity*. Berlin, Germany: Mouton de Gruyter.

Dolman, M., & Spring, R. (2014). To what extent does musical aptitude influence foreign language pronunciation skills? A multi-factorial analysis of Japanese learners of English. *World Journal of English Language, 4*(4), 1–11. https://doi.org/10.5430/wjel.v4n4p1

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching, 31*(3), 117. https://doi.org/10.1017/S026144480001315X

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*(1), 42–68.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. London: Routledge.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In M. Long & C. J. Doughty (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology, 49*(1), 234–250.

Fonseca-Mora, C., Toscano-Fuentes, C., & Wermke, K. (2011). Melodies that help: The relation between language aptitude and musical intelligence. *Anglistik International Journal of English Studies, 22*(1), 101–118.

Gagné, F. (1995). From giftedness to talent: A developmental model and its impact on the language of the field. *Roeper Review, 18*(2), 103–111.

Gagné, F. (2002). *A differentiated model of giftedness and talent* (*DMGT*) (*Update*). Retrieved from http://www.curriculumsupport.education.nsw.gov.au/policies/gats/assets/pdf/poldmgt2002rtcl.pdf.

Gagné, F. (2005). From gifts to talents. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 98–119). New York: Cambridge University Press.

Ganschow, L., & Sparks, R. (1995). Effects of direct instruction in Spanish phonology on the native-language skills and foreign-language aptitude of at-risk foreign-language learners. *Journal of Learning Disabilities, 28*(2), 107–120.

Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 8*(4), 439–454.

Gathercole, S. E., Service E, Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*(1), 65–77.

Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex, 17*(3), 575–582. https://doi.org/10.1093/cercor/bhk001

Golestani, N., & Pallier, C. (2006). Anatomical correlates of foreign speech sound production. *Cerebral Cortex, 17*(4), 929–934.

Golestani, N., Paus, T., & Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron, 35*(5), 997–1010.

Golestani, N., Price, C. J., & Scott, S. K. (2011). Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *Journal of Neuroscience, 31*(11), 4213–4220. https://doi.org/10.1523/JNEUROSCI.3891-10.2011

Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language, 109*(2), 55–67.

Gordon, E. E. (1980). The assessment of music aptitudes of very young children. *The Gifted Child Quarterly, 24*(3), 107–111.

Gordon, E. E. (2001). *Music aptitude and related tests*. Chicago: GIA Publications.

Granena, G., & Long, M. (Eds.). (2013). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: John Benjamins.

Hackett, T. A. (2009). The evolution of the primate and human auditory system. In J. H. Haas (Ed.), *Evolutionary neuroscience* (pp. 893–905). Oxford, UK: Academic.

Hackett, T. A. (2015). Anatomic organization of the auditory cortex. *Handbook of Clinical Neurology, 129*, 27–53. https://doi.org/10.1016/B978-0-444-62630-1.00002-0

Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., & Reiterer, S. M. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: Behavioral predictors and neural substrates. *Brain and Language, 127*(3), 366–376.

Hulshoff, H. E., Schnack, H. G., Posthuma, D., Mandl, R. C. W., Baaré, W. F., van Oel, C., & Kahn, R. S. (2006). Genetic contributions to human brain morphology and intelligence. *The Journal of Neuroscience, 26*(40), 10235–10242. https://doi.org/10.1523/JNEUROSCI.1312-06.2006

Jilka, M. (2009). Talent and proficiency in language. In G. Dogil & S. M. Reiterer (Eds.), *Language talent and brain activity* (pp. 1–16). Berlin, Germany: Mouton de Gruyter.

Kemmerer, D. (2015). *Cognitive neuroscience of language*. New York: Psychology Press.
Kepinska, O. (2017). *The neurobiology of individual differences in grammar learning*. Dissertation. Leiden: University of Leiden.
Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. O. (2017a). On neural correlates of individual differences in novel grammar learning: An fMRI study. *Neuropsychologia, 98*, 156–168.
Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. O. (2017b). Whole-brain functional connectivity during acquisition of novel grammar: Distinct functional networks de-pend on language learning abilities. *Behavioural Brain Research, 320*, 333–346.
Klein, E. C. (1995). Second versus third language acquisition: Is there a difference? *Language Learning, 45*(3), 419–466.
Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology, 15*(2), 207–212.
Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition, 11*(2), 261–271. https://doi.org/10.1017/S1366728908003416
Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews. Neuroscience, 11*(8), 599–605. https://doi.org/10.1038/nrn2882
Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., & Monzalvo, K. (2015). New human-specific brain landmark: The depth asymmetry of superior temporal sulcus. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 1208. https://doi.org/10.1073/pnas.1412389112
Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics, 36*(3). https://doi.org/10.1093/applin/amu054
Li, S. (2016). The construct validity of language aptitude. *Studies in Second Language Acquisition, 38*(4). https://doi.org/10.1017/S027226311500042X
Linck, J. A., Hughes, M. M., Campbell, S., Silbert, N. H., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning, 63*(3), 530–566. https://doi.org/10.1111/lang.12011
Meara, P. (2005). *LLAMA language aptitude tests: The manual*. Swansea, UK: Lognostics.
Milovanov, R. (2009). The connectivity of musical aptitude and foreign language learning skills: Neural and behavioural evidence. *Anglicana Turkuensia, 27*(27), 1–56.
Milovanov, R., Huotilainen, M., Esquef, P. A. A., Välimäki, V., Alku, P., & Tervaniemi, M. (2009). The role of musical aptitude and language skills in preattentive duration determination in school-aged children. *Neuroscience Letters, 460*, 161–165.
Milovanov, R., Huotilainen, M., Velimäki, V., Esquef, P. A. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research, 1194*, 81–89.
Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. A. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences, 20*, 56–60.
Milovanov, R., & Tervaniemi, M. (2011). The interplay between musical and linguistic aptitudes: A review. *Frontiers in Psychology, 2*. https://doi.org/10.3389/fpsyg.2011.00321
Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364).
Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research, 308*, 84–97. https://doi.org/10.1016/j.heares.2013.09.012
Nardo, D., & Reiterer, S. M. (2009). Musicality and phonetic language aptitude. In G. Dogil & S. M. Reiterer (Eds.), *Language talent and brain activity* (pp. 213–256). Berlin, Germany: Mouton de Gruyter.
Neubauer, A. C., & Fink, A. (2009). Intelligence and neural efficiency: Measures of brain activation versus measures of functional connectivity in the brain. *Intelligence, 37*(2), 223–229.

Patel, D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology, 2*(142). https://doi.org/10.3389/fpsyg.2011.00142

Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences, 1191*(1), 62–88.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage, 62*(2), 816–847.

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A. S., McNamara, J. O., & Williams, S. M. (2001). *Neuroscience*. Sunderland, MA: Sinauer Associates.

Reiterer, S. M., Berger, M. L., Hemmelmann, C., & Rappelsberger, P. (2005). Decreased EEG coherence between prefrontal electrodes: A correlate of high language proficiency? *Experimental Brain Research, 163*, 109–113.

Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., et al. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology, 2*(271). https://doi.org/10.3389/fpsyg.2011.00271

Richardson, F. M., & Price, C. J. (2009). Structural MRI studies of language function in the undamaged brain. *Brain Structure and Function, 213*(6), 511–523.

Robinson, P. (2002). Effects of individual differences in intelligence, aptitude, and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfield and Hernstad (1991). In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211–265). Amsterdam: John Benjamins.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73.

Robinson, P. (2012). Individual differences, aptitude complexes, SLA processes, and aptitude test development. In *New perspectives on individual differences in language learning and teaching* (pp. 57–75). Berlin/Heidelberg: Springer.

Rogers, V. E., Meara, P., Aspinall, R., Fallon, L., Goss, T., Keey, E., & Thomas, R. (2016). Testing aptitude. *EUROSLA Yearbook, 16*(1), 179–210.

Rota, G., & Reiterer, S. M. (2009). Cognitive aspects of pronunciation talent: How empathy, mental flexibility, working memory and intelligence interact with phonetic talent. In G. Dogil & S. M. Reiterer (Eds.), *Language talent and brain activity* (pp. 67–96). Berlin, Germany: Mouton de Gruyter.

Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *IRAL – International Review of Applied Linguistics in Language Teaching, 46*(2). https://doi.org/10.1515/IRAL.2008.005

Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319–353). Cambridge, UK: Cambridge University Press.

Schneider, P., Andermann, M., Wengenroth, M., Goebel, R., Flor, H., Rupp, A., & Diesch, E. (2009). Reduced volume of Heschl's gyrus in tinnitus. *NeuroImage, 45*, 927–939.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience, 5*(7), 688–694.

Schneider, P., Sluming, V., Roberts, N., Bleeck, S., & Rupp, A. (2005). Structural, functional, and perceptual differences in Heschl's gyrus and musical instrument preference. *Annals of the New York Academy of Sciences, 1060*(1), 387–394. https://doi.org/10.1196/annals.1360.033

Schneiderman, E. I., & Desmarais, C. (1988a). The talented language learner: Some preliminary findings. *Second Language Research, 4*(2), 91–109. https://doi.org/10.1177/026765838800400201

Schneiderman, E. I., & Desmarais, C. (1988b). A neuropsychological substrate for talent in second-language acquisition. In L. K. Obler & D. Fein (Eds.), *The exceptional brain. Neuropsychology of talent and special abilities* (pp. 103–126). New York: The Guilford Press.

Schumann, J. H. (2004). The neurobiology of aptitude. In J. Schumann et al. (Eds.), *The neurobiology of learning: Perspectives from second language acquisition* (pp. 7–22). Mahwah, NJ: Erlbaum.

Seither-Preisler, A., Parncutt, R., & Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy, and attentional skills in children. *The Journal of Neuroscience, 34*, 10937–10949. https://doi.org/10.1523/JNEUROSCI.5315-13.2014

Serrallach, B., Gross, C., Bernhofs, V., Engelmann, D., Benner, J., Gündert, N., & Seither-Preisler, A. (2016). Neural biomarkers for dyslexia, ADHD and ADD in the auditory cortex of children. *Frontiers in Neuroscience, 10*, 324. https://doi.org/10.3389/FNINS.2016.00324

Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute? *Studies in Second Language Learning and Teaching, 7*(1), 89–103. https://doi.org/10.14746/ssllt.2017.7.1.5

Skehan, P. (1986). Where does language aptitude come from? In P. M. Meara (Ed.), *Spoken language* (pp. 95–113). London: Centre for Information and Language Teaching: British Association of Applied Linguistics.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–95). Amsterdam: John Benjamins.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science, 17*(8), 675–681.

Sparks, R., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics, 21*, 90–111.

Sparks, R. L., Humbach, N., Patton, J. O. N., & Ganschow, L. (2011). Subcomponents of second-language aptitude and second-language proficiency. *The Modern Language Journal, 95*(2), 253–273.

Spolsky, B. (1995). Prognostication and language aptitude testing, 1925-62. *Language Testing, 12*(3), 321–340.

Stansfield, C. W., & Reed, D. J. (2004). The story behind the Modern Language Aptitude Test: An interview with John B. Carroll (1916–2003). *Language Assessment Quarterly: An International Journal, 1*(1), 43–56.

Stern, E., & Neubauer, A. (2013). *Lernen macht intelligent: Warum Begabung gefördert werden muss*. München, Germany: DVA.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: An approach to cerebral imaging*. New York: Thieme.

Thompson, A. S. (2013). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *The Modern Language Journal, 97*(3), 685–701.

Van den Noort, M. W., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11*(4), 289–296.

Vangehuchten, L., Verhoeven, V., & Thys, P. (2015). Pronunciation proficiency and musical aptitude in Spanish as a foreign language: Results of an experimental research project. *Revista de Lingüística y Lenguas Aplicadas, 10*(1), 90–100. https://doi.org/10.4995/rlyla.2015.3372

Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., & Kraus, N. (2012). Relating structure to function: Heschl's Gyrus and acoustic processing. *Journal of Neuroscience, 29*(1), 61–69. https://doi.org/10.1523/JNEUROSCI.3489-08.2009

Wen, Z. (2012). Foreign language aptitude. *ELT Journal, 66*(2), 233–235. https://doi.org/10.1093/elt/ccr068

Wen, Z. (2016). *Working memory and second language learning: Towards an integrated approach*. Bristol, UK: Multilingual Matters.

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: Building and supporting a case for "working memory as language aptitude". *Ilha do Desterro: A Journal of English Language, Literatures in English and Cultural Studies, 60*(60). https://doi.org/10.5007/2175-8026.2011n60p015

Wen, Z., Skehan, P., & Biedroń, A. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching, 50*(1), 1–31.

Winke, P. (2013). An investigation into second language aptitude for advanced Chinese language learning. *The Modern Language Journal, 97*(1), 109–130.

Wong, P. C. M., Perrachione, T. K., & Parish, T. B. (2007). Neural characteristics of successful and less successful speech and word learning in adults. *Human Brain Mapping, 28*, 995–1006.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Science, 6*(1), 37–46.

Let the Music Speak: Examining the Relationship Between Music and Language Aptitude in Pre-school Children

Markus Christiner

Abstract Research focusing on musical expertise and its relation to language function has gained increasingly more interest in the past decade. Various interdisciplinary investigations focusing on musical expertise and language functions reported positive correlations between both faculties. Evidence, therefore, is mounting that musical abilities (e.g. singing ability and the instrument playing) and working memory capacity are the most important mechanisms for predicting individual differences in imitating, memorizing and repeating unfamiliar (foreign) speech material among adults. Most investigations, however, tested people who were musically trained leading to the fact that educational influence undeniably was impacting the performances of individuals. Thirty-six pre-school children with no training in music and foreign language learning between the ages 5 and 6 were tested for their ability to discriminate paired musical statements (PMMA), their singing ability, their ability to remember strings of numbers and their ability to repeat Turkish a language that was completely unfamiliar to the participants. The results revealed that the participants who performed better in the musicality test also had better results in the imitation tasks and possessed high working memory capacity compared to their peers who scored lower in musicality measurements. It can therefore be concluded that musical expertise and talent for speech imitation are linked in children. It seems to be the case that there are innate factors which predetermine musical expertise and positively affect speech imitation aptitude as well. Moreover, the success rate in language and music acquisition processes can be seen as based on both innate and educational factors.

M. Christiner (✉)
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria
e-mail: markus.christiner@univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_8

1 Introduction

Musical expertise has been shown to have a large effect on language functions and thus especially improves the ability to correctly generate new foreign languages orally (Hu et al., 2013; Milovanov, 2009; Nardo & Reiterer, 2009; Pastuszek-Lipinska, 2008; Reiterer et al., 2011; Schön, Magne & Besson, 2004; Thompson, Schellenberg & Husain, 2004; Wong & Perrachione, 2007) – an ability which is claimed to be less successful in adult language learners. Following traditional theories, language acquisition processes undergo critical periods in which languages are easily taken up, leading to a general assumption: the earlier foreign languages are acquired, the better the proficiencies of speakers (e.g. Moyer, 2014). Comparable to language learning, music acquisition processes may develop in an analogous way as both language and music perception show large overlaps in several brain regions (Koelsch et al., 2009; Patel, 2011; Schulze & Koelsch, 2012; Schulze, Zysset, Mueller, Friederici & Koelsch, 2011; Williamson, Baddeley & Hitch, 2010).

The similarity of music and language acquisition and the fact that both music and speech are acoustic phenomena may explain why various studies have shown that people with higher aptitude in music also show higher language imitation abilities (phonetic aptitude). The reasons for the interconnectivity of musical expertise and speech imitation talent may lie in shared brain regions for music and language (Koelsch et al., 2009; Patel, 2011; Schulze & Koelsch, 2012; Schulze et al., 2011; Williamson et al., 2010), improved perceptual abilities in musicians (Oikkonen et al., 2015), enhanced somatosensory skills in particular in singers (Christiner & Reiterer, 2013, 2015; Kleber, Veit, Birbaumer, Gruzelier & Lotze, 2010), or in an improved phonological short-term memory (Christiner & Reiterer, 2013). The latter is also true for contrasting foreign consonants where participants with an improved phonological short-term memory outperformed their lower-scoring counterparts (Cerviño-Povedano & Mora, 2015). Notably, playing musical instruments or singing may improve the perception and production of new speech material as additional training. More recently, though, studies first suggested gene-related effects leading to the development of an improved auditory pathway which, in turn, leads to higher musical aptitude at least on the basis of music perception (e.g. Oikkonen et al., 2015).

1.1 Working Memory

Studies on adults have shown that working memory ability and attentive skills contribute significantly to aptitude for imitating and repeating unfamiliar speech material (Aliaga-Garcia, Mora & Cerviño-Povedano, 2011; Christiner & Reiterer, 2013, 2015). Working memory, described as a phonological store, holds memory traces for some seconds before they fade (Baddeley, 2003) and has been shown to be a good predictor for the imitation of speech and the discrimination of incongruities in paired musical statements or in singing unfamiliar songs in adults (Christiner & Reiterer, 2013, 2015, 2016; Hu et al., 2013; Reiterer et al., 2011). The underlying mechanism why working memory capacity is a good predictor to explain variances in music and foreign language performance may be reliant on large overlaps of tonal and verbal material in the auditory short term memory (Koelsch et al., 2009; Schulze & Koelsch, 2012; Schulze et al., 2011; Williamson et al., 2010;) which may explain why musical training is a benefit for both improving working memory span and oral language skills (Christiner & Reiterer, 2013). While considerable attention has been paid to working memory in adults and its relation to speech imitation and foreign accent imitation, less is known about the nature of preschool children's working memory capacity and its contribution to the imitation of foreign speech material. What has been revealed so far is that working memory capacity affects reading skills and achievements in mathematics in children (e.g. Fitzpatrick & Pagani, 2012), to name but a few. Recent research has also demonstrated that the working memory capacity of children is related to musical abilities (Strait, Hornickel & Kraus, 2011) and to the manipulation and maintenance of information in general (Buschkuehl, Jaeggi & Jonides, 2012). Thus investigations in WM and its relation to educational progress isolated that WM deficits have had an enormous impact on children's achievements (Gathercole, Alloway, Willis & Adams, 2006). Developmental studies found frontal, parietal and striatal brain regions to be most related to WM functions (Bunge & Wright, 2007). Working memory is age-related and most important to acquire new skills and, for example, new language material (Loosli, Buschkuehl, Perrig & Jaeggi, 2012). WM training seems to be highly important for children in general. Performance improvements after training have been observed among children with attention deficits and not only in the tasks trained but also in non-trained WM tasks (Klingberg, Forssberg & Westerberg, 2002). Testing working memory ability of preschool children, however, is a challenging task as they are known to have difficulties remaining focused and managing cognitive loads. Studies have shown that children at around 5 years can manage to memorize around three digits of strings of numbers in a forward order and around two words in a forward order recalled (Roman, Pisoni & Kronenberger, 2014).

1.2 First Language Acquisition and Music Acquisition Processes

Similar to language learning, infants acquire the music which typically represents their own culture, including culture-specific elements. Language and music acquisition share basic aspects and influence each other (Patel & Daniele, 2003), but both faculties differ in how they are arranged and organised (McMullen & Saffran, 2004). From a developmental point of view, it has been argued that infants have prenatal experience in music and language and show preferences for their culture-specific music and their first language shortly after birth (De Casper & Fifer, 1980; McMullen

& Saffran, 2004). Infants, for instance, are capable of distinguishing and acquiring all different types of phonemes in the first 6 months of their lives (Kuhl, 2004), an ability which gets largely lost at later stages as a consequence of mother tongue specification (Patel, 2008). Music acquisition processes show similar aspects, even though they are not as obvious as is the case with language acquisition. One-year-old infants who are exposed to particular, yet unknown music show sensitivity towards the new musical rhythm after a short period of time (Hannon & Trehub, 2005), while another study found that, in marked contrast to infants, adults could not detect non-isochronous metres of Balkan music, a type of music which was unfamiliar to American participants showing an analogy to being exposed to foreign languages for the first time (Hannon & Trainor, 2007). This in particular shows that there are music acquisition processes which can be compared to language acquisition processes, especially to second language acquisition, despite it being considered less successful and more difficult during adulthood.

Turning to music acquisition processes, researchers have argued that music instruction cannot speed up children's performance on specific elements such as on harmonic perception, which may be more of a natural development (e.g. Costa-Giomi, 2003). What can be improved in children, however, are those elements which are already acquired, such as recognition of culture specific rhythm (ibid., 2003). Gordon (2003), for instance, argues that infants are born with a certain musical aptitude and the musical expertise someone reaches is the result of innate and educational factors, i.e. the earlier infants are exposed to an environment rich in music, the better they may be later as musical aptitude has been said to decrease immediately after birth. The ability to give similar judgements like adults about music seems to be developed at the age of nine (Sloboda, 2005) showing that children have acquired communicative competence significantly earlier.

Taking into account comparative music and language studies, investigations included participants which were trained either in languages or in playing musical instruments which cannot inform about their potential without being influenced by educational differences.

Pre-school children would be ideal test candidates for doing aptitude research for music and foreign accent imitation for three reasons. Firstly, they hardly ever train musical instruments or receive singing education in their first years of life. Secondly, their musical ability may be largely natural and unshaped in terms of educational influence. Thirdly, they do not get formal instructions for language learning or participate in foreign language learning programmes at all. Consequently, individual differences may be based on their aptitude rather than on training received during this particular life-span. If children with higher musical aptitude are also better in accent imitation it may be that first language acquisition and music acquisition are ultimately linked processes united under an aptitude which may globally be responsible for the potential of the acquisition of any acoustic parameter.

1.3 Studies on the Relationship Between Musical Expertise and Speech Imitation

Considering current research, comparative speech and music studies most often contrasted the speech imitation performances of professional musicians and non-musicians which predetermined research design as in a professional context adult participants are available only (e.g., Christiner & Reiterer, 2013, 2015; Nardo & Reiterer, 2009). As far as known, researchers who worked on musical skills in children most often included musical training aiming at demonstrating whether teaching methods, which include musical input, contribute to higher language abilities or not, concluding that musical input improves verbal abilities as well (Fonseca-Mora, Jara-Jiménez & Gómez-Domínguez, 2015; Moreno et al., 2009, 2011). Other studies were conducted with children who had musical experience to some extent (Milovanov, 2009).What all these studies, however, have in common is that musical expertise goes hand in hand with higher language imitation abilities suggesting that musical input/training could be integrated in foreign language teaching (e.g., Christiner & Reiterer, 2013, 2015; Fonseca-Mora et al., 2015). Another commonality of the former studies mentioned is that some participants received musical training while testing and were influenced this training. This, however, also raises the question whether musical education is responsible for individual differences in performances, or whether there may be a natural aptitude which predetermines the potential of people's verbal and musical ability.

I set out to test this notion and selected kindergarten children (N = 36) who had never received musical education, i.e. no individual musical training or instrumental lessons, apart from traditional singing activities in kindergarten. Their ability to discriminate rhythmical or tonal changes in paired musical statements, as well as their ability to repeat strings of numbers, and their ability to imitate foreign speech material in Turkish was tested. I aimed at isolating whether earlier results on adults, which revealed a significant relationship between musical expertise, working memory capacity and singing to speech imitation would yield similar results by testing pre-school children who did neither receive foreign language education nor musical training.

RQ1. If musical abilities were also related to children's speech imitation, it could be further evidence that phonetic and musical aptitude are highly intertwined by nature and not a result of training only.

RQ2. Is working memory also one of the best predictors for speech imitation in pre-school children?

2 Materials and Methods

2.1 Participants

For this study, I selected 36 pre-school children between the ages of 5 and 6 years with no prior experience in Turkish and no instrumental lessons taken before the onset of the study. During that time the participants had limited knowledge in singing apart from singing traditional songs and were able to count at least to ten. All participants were German native speakers and had no experience in foreign languages and just limited knowledge of very basic words in English. The participants' parents gave informed consent to the study and reported that the children had no hearing impairments or any other nuisance.

2.2 Behavioural Testing: Turkish Imitation

In the behavioural testing I analysed the participants' ability to repeat foreign language material in Turkish. The original phrases were recorded in a soundproof room and spoken by native speakers for each language. The participants were instructed to listen to the phrases and were ordered to repeat what they had heard in the best way possible after listening to them tree times. The length of the language material was five syllables since the WM capacity of children is still developmental and thus limited compared to adults (phrases see appendix). For adult testing, language material of around 9–11 syllables, depending on the typology, is appropriate (see: Christiner, Rüdegger & Reiterer, 2017). For simulating the procedure we used three different phrases before the real testing took place and thus the children knew the entire procedure. The average time of testing and preparing was around 5 min.

For rating the sound files of the participants we instructed seven native speakers to evaluate how native-like the participants' performances were. The raters had to indicate their response on a scale of 0–10. Zero was the lowest score someone could receive, while ten was the highest score for a native-like performance.

2.3 Behavioural Testing: Musicality Test (PMMA), Singing Ability and Working Memory Tests

As a measurement to explore the participants' musical expertise, the PMMA test (Primary Measures of Music Audiation, Gordon, 2006) was used. This test was invented to test children in grades from kindergarten to third grade and it elicits children's ability to discriminate tonal and rhythmical changes of paired musical statements. The children's singing ability was rated by the caretakers of the kindergarten, who rated two different aspects on a scale from 0 (lowest score) to 10 (highest score). One aspect rated was how well the children sang and the second aspect was how often the children started to sing without being stimulated to sing. Furthermore, over a period of 2 weeks, the parents were asked to report how many hours on average the children sang a day. For testing the children's acoustic working memory we used a simplified version of the working memory test based on the Wechsler test for adults (1939). The modified working memory test consists of a digit span forward only and excludes the backwards part. The participants had to repeat strings of numbers in German. Furthermore, we also aimed at testing the recall of non-words of the phonetic quality in German and used a non-word span in German (Benner, 2005).

2.4 Questionnaire

For further analysis the participants' parents were asked to fill out an additional questionnaire. This questionnaire aimed at eliciting background knowledge of the participants' surroundings and environment and whether the participants had received or were receiving any training which could influence the results of our assumptions.

3 Results

For the validity of the data, crucial factors such as the participants' age, gender, and handedness were evaluated. The data showed a normal distribution for all parameters. The speech stimuli analysed consisted of three phrases of either five syllables. All three phrases together comprise the Turkish performance (Turkish imitation) which was taken for further analysis. The mean of the Turkish imitation score of the preschool children was 3.03, $SD = 1.74$. The lowest score was .3 and the highest score was 7.09. The mean of the working memory test, the digit span forward in German, was 3.7, $SD = 1.19$. The maximum score was 8 and the lowest score 2. The mean of the working memory test, the digit span forward in English, was 2.42, $SD = 1.26$. The lowest score was 2, while the highest score was 6. The mean of the non-words working memory test was 2.78, $SD = 0.96$. The highest score which was achieved was 6 and the lowest 1. For the musical abilities we used the PMMA musicality test. The mean of the discrimination task where the children had to detect changes in melody in paired musical statements, the tonal parameter, was 28.5, $SD = 4.25$. The highest score was 36 and the lowest score 18. The mean of the second musicality task where participants had to detect rhythmical changes in paired melodies was 24.64, $SD = 3.76$. The mean of the total score of the musicality test was 53.11, $SD = 6.52$. The highest score was 65 and the lowest score 38. The mean of the singing score of how well the children sang was 6.47, $SD = 0.70$ and the lowest score was .25, whereas the highest score was 10 (Tables 1 and 2).

Table 1 Mean, standard deviation, the minimum and the maximum value of the most important variables discussed

	M	SD	Min	Max
Working memory F	3.7	1.19	2.00	8.00
Non-words WM	2.78	0.96	1.00	6.00
PMMA tonal	28.50	4.25	18.00	36.00
PMMA rhythm	24.64	3.76	17.00	34.00
Singing ability	6.47	0.70	0.25	10.00
Singing intuitive	5.45	2.42	0.25	10.00
Turkish performance	3.03	1.74	0.3	7.09

3.1 Results of the Turkish Imitation Task

Results revealed that the working memory test forward (WM) was significantly related to how well the participants imitated Turkish $r_s = 0.57$, p (one-tailed) < 0.01. Furthermore, Turkish was significantly correlated with the non-words working memory test $r_s = 0.31$, p (one-tailed) < 0.05. The Turkish performances also showed a significant relationship to the musicality test (PMMA). The tonal parameter of the musicality test correlated significantly with the Turkish performance $r_s = 0.33$, p (one-tailed) < 0.05. The total score of the musicality test (PMMA) also showed correlations to the Turkish performances, $r_s = 0.38$, p (one-tailed) < 0.05. Interestingly, the parameter rhythm of the musicality test showed no effect on the imitation ability of Turkish. The Turkish imitation, however, also showed a positive correlation with intuitive singing $r_s = 0.35$, p (one-tailed) < 0.05.

3.2 Working Memory Test: Forward Digit Span

Like for adults, the ability to repeat numbers shows high correlations with the Turkish imitation ability of children, $r_s = 0.57$, p (one-tailed) < 0.01. The non-words working memory test was also significantly related to the working memory test (forward digit), $r_s = 0.52$, p (one-tailed) < 0.01. The working memory test forward was also correlated to the musicality test (PMMA) of the total score, $r_s = 0.53$, p (one-tailed) < 0.01 and the tonal score $r_s = 0.57$, p (one-tailed) < 0.01.

3.3 Musicality Test (PMMA)

The total score of the musicality test correlated significantly with the working memory test forward digit $r_s = 0.53$, p (one-tailed) < 0.01 and the non-words memory test $r_s = 0.51$, p (one-tailed) < 0.01. There was also a significant relationship between the tonal musicality parameters and the working memory test (forward digit) $r_s = 0.57$,

Table 2 (N = 36) Correlations of the individual tasks, working memory test, the musical measurements and the foreign language imitated (Turkish)

Correlations (Spearman)

		Working memory test forward	Non-words working memory	PMMA tonal	PMMA rhythm	PMMA total	Singing score	Intuitive singing	Turkish performance	Number of hours caretakers sing with the child	Number of hours the child sings
Working memory test forward	*rs*	1	.52**	.57**	.25	.53**	.34*	.30*	.57**	.12	.09
Non-words WM	*rs*	.52**	1	.42**	.32*	.51**	.30*	.21	.31*	.17	−.08
PMMA tonal	*rs*	.57**	.42**	1	.30*	.81**	.38*	.23	.33*	.10	.09
PMMA rhythm	*rs*	.25	.32*	.30*	1	.76**	.39**	.39**	.25	.22	.42**
PMMA total	*rs*	.53**	.51**	.81**	.76**	1	.47**	.34*	.38*	.25	.26
Singing score	*rs*	.34*	.30*	.38*	.39**	.47**	1	.81**	.20	.24	.36*
Intuitive singing	*rs*	.30*	.21	.23	.39**	.34*	.81**	1	.35*	.17	.37*
Turkish performance	*rs*	.57**	.31*	.33*	.25	.38*	.20	.35*	1	−.09	.19
Number of hours caretakers sing with the child	*rs*	.12	.17	.10	.22	.25	.24	.17	−.09	1	.43**
Number of hours the child sings	*rs*	.09	−.08	.09	.42**	.26	.36*	.37*	.19	.43**	1

**Correlation is significant at the 0.01 level (1-tailed)
*Correlation is significant at the 0.05 level (1-tailed)

p (one-tailed) < 0.01 and the non-words working memory test $r_s = 0.42$, p (one-tailed) < 0.01. The tonal musicality parameter also correlated with the Turkish performances, $r_s = 0.33$, p (one-tailed) < 0.05. Furthermore, the rhythmic parameter of the musicality test correlated significantly with the hours children were singing $r_s = 0.42$, p (one-tailed) < 0.01.

3.4 Singing Ability

How well the children sang correlated with the working memory digit span forward $r_s = 0.34$, p (one-tailed) < 0.05 and non-word span $r_s = 0.30$, p (one-tailed) < 0.05. In addition, how well they sang correlated significantly with the musicality measurements. There was a significant correlation with the tonal discrimination task $r_s = 0.38$, p (one-tailed) < 0.05, the rhythmic discrimination ability $r_s = 0.39$, p (one-tailed) < 0.01 and the total score of the musicality test $r_s = 0.47$, p (one-tailed) < 0.01. The intuitive singing score, which informs about how often they sang without being asked to sing, shows correlations with the rhythmic music discrimination task $r_s = 0.39$, p (one-tailed) < 0.01 and the Turkish performance of the participants $r_s = 0.35$, p (one-tailed) < 0.05. The number of hours a child sang correlated with the rhythmic musicality measurements $r_s = 0.42$, p (one-tailed) < 0.01. The number of hours the caretaker sang or played musical instruments for the children did not contribute to their achievement in any cognitive skills investigated.

3.5 Dependent T-Test PMMA Tonal Versus Rhythmical Discrimination Ability

For isolating whether there are mean differences in the musical perception tasks (PMMA tonal and rhythm) a dependent t-test was performed. The results have revealed that on average tonal discrimination ability was significantly more developed in pre-school children (M = 28.56, SE = 4.21) than the rhythmical discrimination ability (M = 24.58, SE = 3.79), $t(35) = 5.12$, $p < 0.01$, $r = 0.64$.

3.6 Independent T-Test Singing Ability and Music Perception (Rhythm)

For analysing singing ability and its influence on music perception an independent t-test was performed. Grouping of high and low singing ability was based on the median of the singing ratings (Mdn = 6.5). Results have shown that on average better singers (M = 25.89, SE = 2.76) were also better in the rhythmic music perception task (PMMA rhythm) than the lower singing ability group ($M = 23.2941$, $SE = 4.41$). This difference was significant effect $t(33) = 2.10$, $p < 0.05$; which did represent a medium-sized effect $r = 0.34$.

4 Discussion

The results of this investigation revealed that the relations between speech imitation, working memory and musical expertise are similar in children and in adults. Previous investigations on children have shown that musical expertise leads to higher language abilities (Milovanov, 2009). There are studies that incorporated musical training which demonstrated that long-term and short-term musical training improves verbal and musical abilities in adults and children (e.g. Bangert et al., 2006; Moreno et al., 2009). Receiving musical education also seems to have an effect on brain plasticity in children in a very short period of time (Moreno et al., 2011) leaving no doubt that a rich musical environment improves the control of attention and memory affecting verbal and musical ability. Both music and language systematically develop through training, which in turn, questions whether musical expertise also shows a high correlation to verbal abilities in children that did not receive education or participated in any musical training before. This study, therefore, included pre-school children that did not receive formal music education aiming at uncovering whether individual differences can even be observed among children that were naïve in that domain. If so, it would allow us to conclude that individual differences are based on their aptitude. Indeed, results showed that musical expertise and speech imitation ability also correlated significantly in children who are untrained which suggests that there is a natural aptitude for acoustic parameters irrespective of whether signals belong to speech or music. This would mean that the relation between both faculties is not merely a result of education, development and exposure, but possibly also by nature. Even though the test design of this study does not allow insights in genetic parameters and inherent abilities to be given, it may be suggested that the genetic constitution of children or inherent factors may be responsible for individual differences in the performances of speech imitation abilities and musical expertise in early childhood.

Interdisciplinary studies on music and language have shown that there is a significant relationship between the ability to repeat new and unfamiliar speech material and their ability to repeat strings of numbers and to discriminate musical statements in adult musicians, that is to say instrumentalists and vocalists (Christiner & Reiterer, 2013, 2015). The same effect, even though slightly less dominant, has been found in this research where children with higher aptitude for imitating Turkish were also better at detecting tonal changes/incongruities of musical statements as well as in the overall performance of the musicality test (PMMA). This suggests that musical expertise and language functions are highly intertwined from early childhood on. Considering that the children of this investigation were naïve and had received no education in music or in foreign languages, their performances were based on individual differences gained outside of the formal educational context. Research on musical aptitude claim to have first evidence that the genetic constitution of children influences talent for music (Oikkonen et al., 2015). Participants of the latter study showed alterations in the development of the auditory pathway that influence the way auditory signals are processed leading to a higher ability to discriminate incongruities in melodies or pitches (ibid., 2015). This, however, does not mean that musical training is not as effective as genetic benefits may be, but it shows that genetic factors may play an important role in facilitating and manifesting aptitudes which may explain individual differences and different success rates in language acquisition processes among peers that receive the same education.

Referring to the field of psychology, aptitude has been described as an untrained natural ability evident in at least one ability domain. This ability places individuals at least among the upper 15% of their peers (Gagné, 2005). Psychologists argue that gifts transform into talents through systematic training (ibid., 2005). In the field of music, it has been suggested that musical talent is reliant on both innate potential and early influences (Gordon, 2003). Interestingly, the results of this investigation have shown that the rhythmic parameter of the musicality test did not show any correlation to the ability to repeat Turkish showing that tonal and rhythmical abilities seem to develop differently with the tonal ability developing first. On a developmental basis it has been claimed that children have similar musical abilities to adults at around the age of nine (ibid., 2003) which demonstrates that communicative competence develops earlier. Considering the results of this study, it can be suggested that the higher scoring in the tonal ability of children between the ages 5 and 6 shows that tonal discrimination ability may be developed earlier than the ability to discriminate rhythmical changes in tonal statements. It could be speculated that this may be related to early language acquisition processes and reliant on the nature of language learning of infants which is more musical in its nature. Ontogenetically, the vocalisation of infants resembles more singing than speech and incorporates more tonal aspects than adult speech (Murphey, 1990). Infants generate speech patterns without any concern and focus on meaning but purely out of pleasure in a song-like manner (Murphey, 1990). This suggests two aspects. One is that the segmentation of individual phonemes and the production of isolated words may be not possible for infants because they lack the ability to generate precise speech patterns immediately. Thus they practice tonal and melody-like utterances first. Another is that they lack the ability to perceive rhythmical aspects of the language they learn. However, the latter is very unlikely as some studies have showed that infants exposed to unfamiliar music show sensitivity to new musical rhythms after a short period of time (Hannon & Trehub, 2005). Infants' inability to produce speech is more reliant on motor control. First language acquisition develops alongside motor control whereby infants continually expand their oral language performance by getting control over their vocal apparatus (Iverson, 2010). This way to get control over the body induces/alters brain structure in infants. What has been observed in infants seems to be completely reasonable, but not for adults. However, to a certain extent, the same effect can be noted in adults who continually work on their vocal ability. Vocalists, for example, show alterations and structural adaptations in the dorsal and ventral regions of the brain which have not yet been detected in instrumentalists (Halwani, Loui, Rüber, & Schlaug, 2011). Those alterations seem to be induced by long-term vocal motor training. Behavioural research has also shown that vocalists outperform instrumentalists in language tasks (Christiner & Reiterer, 2015). To turn back to the findings of this investigation, singing ability did not contribute to speech imitation in this investigation, however, singing ability, intuitive singing behaviour and the hours the children sang were significantly related to children's ability to discriminate rhythmical differences in paired musical statements (see independent t-test Sect. 3.6) which may illustrate that singing improves musical abilities and maybe in turn language functions. Additionally, intuitive singing behaviour also correlated with Turkish imitation which questions in how far intrinsic motivation to vocal behaviour is influencing language talent in general. Interestingly, the number of hours the caretakers sang with their children did not contribute to any of the variables in consideration apart from how often the children sang. Singing ability is still difficult to be analysed and for future research it is necessary to find further ways of measuring singing ability accurately, apart from ratings to come closer to the conundrum of explaining the link between song, music and language functions.

To recap, the working memory test showed high correlations with the imitation of Turkish and the ability to discriminate incongruities in musical statements. Working memory assists multiple cognitive processes for the storing and the manipulation of information (Loosli et al., 2012). Individual differences in the ability to remember chunks or items are especially age-related (ibid., 2012). Children are said to be able to remember around three digits of strings of numbers in a forward order and around two words in a forward order recalled (Roman et al., 2014). The study's findings are in line with previous research, not only in the items and chunks remembered by the participants, but also in terms of WM's link to language functions (Juffs & Harrington, 2011; Linck, Osthus, Koeth & Bunting, 2014; Rota & Reiterer, 2009). People who are in the possession of high working memory capacity are also excellent at processing foreign language material (Skehan, 1998). As in this investigation, WM was related to Turkish imitation, other studies additionally found higher working memory in reading (Fitzpatrick & Pagani, 2012), in language comprehension (Daneman & Merikle, 1996) or in mathematics (Fitzpatrick & Pagani, 2012). The pre-school children in this investigation did not receive any WM training; the latter, however, included in the teaching environment might be important for expanding the capacity of items to be remembered in young children whose WM capacity is limited and still developmental compared to adults. Literature focusing on WM training and its effect on other cognitive functions detected changes in executive control (Jaeggi, Buschkuehl, Jonides & Shah, 2011) and remarkable improvements in many cognitive processes (Loosli et al., 2012).

As the WM is concerned with the rapid and temporary storing of information, overlaps between tonal and verbal material are plausible. The tonal discrimination ability and the total performance of the musicality test correlated with the working memory capacity of the participants. Strictly speaking, the imitating of new meaningless language material is quite similar to the learning of new music material which does not carry linguistic meaning. Brain researchers (e.g., Koelsch et al., 2009; Schulze et al., 2011) showed that overlaps of tonal and verbal material may be reliant on working memory capacity. Functional plasticity, in general, has also been said to be induced by music (Koelsch et al., 2009), which suggests that working memory capacity, musical abilities and speech/accent imitation ability are ultimately linked.

5 Conclusion and Future Directions

As shown, speech imitation talent, musical aptitude and working memory capacity seem to be linked in pre-school children similarly like in adults. Individual differences in the performances of the children on all tasks show that there are differences which cannot be explained by formal educational aspects as they were naïve in language, music, singing, and working memory training. Of course, the research design does not allow speculation about genetic differences and inherent pre-existence of aptitudes, even though individual differences in the performances of the participants cannot be explained by educational differences. The answer to this will require future work that could investigate in neuronal correlates and aptitude markers and in how far training effect shapes and redesigns brain structure by conducting longitudinal studies. Then aptitude and talent could be observed more reliably and the proportion of aptitude and training defined. Nevertheless, despite the limitations of this study, it can be suggested that an environment rich in music may be beneficial for children's development cognitively.

Appendix

Turkish Language Material

1. Feberbahceli
2. Ortaögretim
3. Oyuncaklarim

Questionnaire

This survey has been conducted by Markus Christiner, to better understand how music and language acquisition processes of pre-school children are linked. Please read every instruction and write your answers. This is no test and so there are no right or wrong answers. In case you don't know how to answer questions please talk to the experimenter who will give you advice. The results of this survey will be used for research purposes only. Thank you very much for your help.

1.	Identification number
2	Date of birth
3	Handedness
4	For the parents:
	Mother tongue(s) of both parents
5	Which mother tongue does your child speak?
6	Does your child play a musical instrument? (Yes or no)
7	Since when has your child started to play a musical instrument? (If not, please ignore the question)
9	How often does your child sing?
	Indicate yes (everyday), no (if your child never sings) or seldom (e.g., once in a week)
10	How many hours does your child sing in a week on average (e.g., 1 h)
11	For the parents
	Do you regularly sing with your children? For example every week? (Indicate in hours e.g., 0.5 h a week)
12	For the parents:
	Do you play a musical instrument?
13	For the parents:
	Do you play a musical instrument? If so, please indicate how many hours you play a week.
14	For the parents:
	Which foreign languages do you speak? Write all languages down in order of proficiency. This means the foreign language you speak well first, followed by the ones you speak less well.

Many thanks for your participation.

References

Aliaga-García, C., Mora, J. C., & Cerviño-Povedano, E. (2011). L2 speech learning in adulthood and phonological short-term memory. *Poznań Studies in Contemporary Linguistics PSiCL, 47*, 1–14. https://doi.org/10.2478/psicl-2011-0002

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*, 829–839. https://doi.org/10.1038/nrn1201

Bangert, M., Peschel, T., Schlaug, G., Drescher, D., Hinrichs, H., Heinze, H.-J., & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage, 30*, 917–926. https://doi.org/10.1016/j.neuroimage.2005.10.044

Benner, U. (2005). *Syllables in speech production: A study of the mental syllabary*. Linguistic Master thesis at the University of Stuttgart, Stuttgart, Germany.

Bunge, S. A., & Wright, S. B. (2007). Neurodevelopmental changes in working memory and cognitive control. *Current Opinion in Neurobiology, 17*, 243–250. https://doi.org/10.1016/j.conb.2007.02.005

Buschkuehl, M., Jaeggi, S. M., & Jonides, J. (2012). Neuronal effects following working memory training. [supplemental material]. *Developmental Cognitive Neuroscience, 2*, 167–179. https://doi.org/10.1016/j.dcn.2011.10.001

Cerviño-Povedano, E., & Mora, J. C. (2015). Spanish EFL learners' identification of /iː–I/ and phonological short-term memory. In *Proceedings of the 32nd International Conference of the Spanish Association of Applied Linguistics (AESLA): Language Industries and Social Change. Procedia – Social and Behavioral Sciences, 173*, 18–23. https://doi.org/10.1016/j.sbspro.2015.02.024.

Christiner, M., & Reiterer, S. M. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology*, *4*(874). https://doi.org/10.3389/fpsyg.2013.00874.

Christiner, M., & Reiterer, S. M. (2015). A mozart is not a pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience, 9*(482). https://doi.org/10.3389/fnhum.2015.00482.

Christiner, M., & Reiterer, S. M. (2016). Music, song and speech: A closer look at the interfaces between musicality, singing and individual differences in phonetic language aptitude. In G. Granema, D. O. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 131–156). Amsterdam: John Benjamis. https://doi.org/10.1075/bpa.3.07chr

Christiner, M., Rüdegger, S., Reiterer, S. M. (2017). Sing Chinese and tap tagalog? Predicting individual differences in musical and phonetic aptitude using language families differing by sound-typology. *International Journal of Multilingualism.* https://doi.org/10.1080/14790718.2018.1424171

Costa-Giomi, E. (2003). Young children's harmonic perception. *Annals of the New York Academy of Sciences, 999*, 477–484. https://doi.org/10.1196/annals.1284.058

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review, 3*, 422–433.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science New Series, 208*, 1174–1176.

Fitzpatrick, C., & Pagani, L. S. (2012). Toddler working memory skills predict kindergarten school readiness. *Intelligence, 40*, 205–212. https://doi.org/10.1016/j.intell.2011.11.007

Fonseca-Mora, C., Jara-Jiménez, P., & Gómez-Domínguez, M. (2015). Musical plus phonological input for young foreign language readers. *Frontiers in Psychology, 6*, 286. https://doi.org/10.3389/fpsyg.2015.00286

Gagné, F. (2005). From gifts to talents: The DMGT as a developmental model. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of Giftedness* (2nd ed., pp. 98–119). Cambridge, UK: Cambridge University Press.

Gathercole, S. E., Alloway, T. P., Willis, C., & Adams, A.-M. (2006). Working memory in children with reading disabilities. *Journal of Experimental Child Psychology, 93*, 265–281. https://doi.org/10.1016/j.jecp.2005.08.003

Gordon, E. E. (2003). *A music learning theory for newborn and young children*. Chicago: GIA.

Gordon, E. E. (2006). *Primary measures of music audiation*. Chicago: GIA.

Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists and non-musicians. *Frontiers in Psychology, 2*(156), 156. https://doi.org/10.3389/fpsyg.2011.00156

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences, 11*, 466–472. https://doi.org/10.1016/j.tics.2007.08.008

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *PNAS, 102*, 12639–12643. https://doi.org/10.1073/pnas.0504254102

Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., & Reiterer, S. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: Behavioural predictors and neural substrates. *Brain & Language, 127*, 366–376. https://doi.org/10.1016/j.bandl.2012.11.006

Iverson, J. M. (2010). Developing language in a developing body: The relationship between motor development and language development. *Journal of Child Language, 37*, 229–261. https://doi.org/10.1017/S0305000909990432

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and longterm benefits of cognitive training. *PNAS, 108*(25), 10081–10086. https://doi.org/10.1073/pnas.1103228108

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching, 44*, 137–166. https://doi.org/10.1017/S0261444810000509

Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., & Lotze, M. (2010). The brain of opera singers: Experience-dependent changes in functional activation. *Cerebral Cortex, 20*, 1144–1152. https://doi.org/10.1093/cercor/bhp177

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24*, 781–791. https://doi.org/10.1076/jcen.24.6.781.8395

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping, 30*, 859–873. https://doi.org/10.1002/hbm.20550

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5*, 831–843. https://doi.org/10.1038/nrn1533

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*, 861–883. https://doi.org/10.3758/s13423-013-0565-2

Loosli, S. V., Buschkuehl, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology, 18*(1), 62–78. https://doi.org/10.1080/09297049.2011.575772

McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception, 21*, 289–311. https://doi.org/10.1525/mp.2004.21.3.289

Milovanov, R. (2009). Musical aptitude and foreign language learning skills: Neural and behavioural evidence about their connections. In *Proceedings of the 7th triennial conference of European Society for the Cognitive Sciences of Music (ESCOM 2009),* 338–342.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, G. E., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science, 22*, 1425–1433. https://doi.org/10.1177/0956797611416999

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, L. S., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex, 19*, 712–723. https://doi.org/10.1093/cercor/bhn120

Moyer, A. (2014). What's age got to do with it? Accounting for individual factors in second language accent. *Studies Second Languages Learning Teacher, 4*, 443–464. https://doi.org/10.14746/ssllt.2014.4.3.4

Murphey, T. (1990). The song stuck in my head phenomenon: A melodic din in the lad? *System, 18*, 53–64.

Nardo, D., & Reiterer, S. (2009). Musicality and phonetic language aptitude. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 213–256). Berlin, Germany: Mouton de Gruyter.

Oikkonen, J., Huang, Y., Onkamo, P., Ukkola-Vuoti, L., Raijas, P., Karma, K., & Vieland, V. J. (2015). A genome-wide linkage and association study of musical aptitude identifies loci containing genes related to inner ear development and neurocognitive functions. *Molecular Psychiatry, 20*, 275–282. https://doi.org/10.1038/mp.2014.8

Pastuszek-Lipinska, B. (2008). Influence of music education on second language acquisition. In *Proceedings of Acoustics '08 Paris,* 5125–5130.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford, UK: Oxford University Press.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The opera hypothesis. *Frontiers in Psychology, 2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*, B35–B45. https://doi.org/10.1016/S0010-0277(02)00187-7

Reiterer, S., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., et al. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology, 2*(271). https://doi.org/10.3389/fpsyg.2011.00271

Roman, A. S., Pisoni, D. B., & Kronenberger, W. G. (2014). Assessment of working memory capacity in preschool children using the missing scan task. *Infant and Child Development, 23*, 575–587. https://doi.org/10.1002/icd.1849

Rota, G., & Reiterer, S. M. (2009). Cognitive aspects of language talent. In G. Dogil & S. M. Reiterer (Eds.), *Language talent and brain activity*. Berlin, Germany: Mouton de Gruyter.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*(3), 341–349. https://doi.org/10.1111/1469-8986.00172.x

Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences, 1252*, 229–236. https://doi.org/10.1111/j.1749-6632.2012.06447.x

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Human Brain Mapping, 32*, 771–783. https://doi.org/10.1002/hbm.21060

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.

Sloboda, J. A. (2005). *Exploring the musical mind: Cognition, emotion, ability, function*. Oxford, UK: Oxford University Press.

Strait, D. L., Hornickel, J., & Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behavioral and Brain Functions, 7*(7), 44. https://doi.org/10.1186/1744-9081-7-44

Thompson, W. F., Schellenberg, G. E., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion, 4*, 46–64. https://doi.org/10.1037/1528-3542.4.1.46

Wechsler, D. (1939). *The Measurement of adult intelligence*. Baltimore: Williams and Wilkins. https://doi.org/10.1037/10020-000

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: Comparing phonological similarity and pitch proximity. *Memory & Cognition, 38*, 163–175. https://doi.org/10.3758/MC.38.2.163

Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics, 28*, 565–585. https://doi.org/10.1017/S0142716407070312

Language Aptitude in Relation to Handedness, Hemispheric Dominance, Cognitive Learning Strategies and Non-verbal IQ: A Combined Quantitative and Qualitative Study

Klara Kager

Abstract Individuals vary greatly in their aptitude for language, a phenomenon especially visible in the diverging degree of proficiency present in second language learners. This study uses a combined quantitative and qualitative approach to test individual differences in language aptitude. It explores the impact of handedness and hemispheric brain dominance on language performance by testing participants' cognitive flexibility in verbal and non-verbal domains. The test battery, consisting of the fifth part of the Modern Language Aptitude Test (MLAT), part F of the LLAMA language aptitude test, and Raven's Progressive Matrices, was administered to 16 right- and 16 left-handed native German speakers (n = 32) studying art, languages, or natural sciences. These tests serve to evaluate the participants' aptitude for vocabulary learning, grammatical inferencing, and abstract reasoning. Additionally, two (non-validated) complementary questionnaires enquired about the participants' preference for either verbal or non-verbal games. The results confirm that handedness does not have any traceable influence on language aptitude, and the groups of art students, language students, and science students did not produce significantly different results. Correlations between scores on the language and reasoning tests indicate that verbal and non-verbal abilities draw on similar mental resources. An additional finding shows that participants opting for non-verbal games scored significantly higher on the language tests than participants who preferred verbal games. These findings lead to the conclusion that handedness and hemispheric dominance have no measurable effect on language performance. Results further suggest that good visuospatial skills can present a considerable advantage in second language learning.

K. Kager (✉)
Department of Linguistics, University of Vienna, Vienna, Austria
e-mail: klara.kager@univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_9

1 Introduction

There is consensus that the acquisition of a second language results in diverging levels of language proficiency. Dörnyei and Skehan (2003) explain this phenomenon by suggesting that "there is a specific talent for learning foreign languages which exhibits considerable variation between individual learners" (p. 590). The exploration of these differences between individual learners and this concept of talent has gained increasing interest in various fields such as cognitive sciences, linguistics, and language teaching. The main reason for this growing attraction is that analysing the nature of language aptitude and the differences in second language learning success can provide a reference scale for learning achievements, as well as an insight into the phenomenon of variation itself (DeKeyser, 2012). Furthermore, a better understanding of the construct of talent can help to predict an individual's success rate in second language learning. Aptitude and motivation in particular have been pinpointed as good indicators of foreign language learning talent (Dörnyei & Skehan, 2003). Other widely researched factors include gender (Bernat & Lloy, 2007; Boyle, 1987), age (Long, 1990; Saville-Troike, 2006), genetics (Briscoe, Chilvers, Baldeweg, & Skuse, 2012; Dediu, 2008) and musicality (Christiner & Reiterer, 2015).

The logical starting point in the discussion on individual differences in language learning success is the neurobiological mechanisms that underlie language abilities and the cognitive strategies involved in language learning. Despite its centrality to numerous scientific questions regarding issues including – but certainly not limited to – language, the brain has taken centre stage in the last 15 years with the advent of neuroimaging techniques and the establishment of new research fields such as cognitive sciences, second language acquisition (henceforth SLA), and neuro- and psycholinguistics (Reiterer, 2009). Today, SLA is a popular and widely researched topic in linguistics, producing an abundance of studies that highlight different aspects important for language acquisition. However, the focus of most studies lies on the product of learning, leaving the process and the neural underpinning of language largely unexplored (DeKeyser, 2012).

The main reason why the neural processes of language acquisition remain uncertain even today is that observing the brain presents a challenging venture. For centuries, the only option to explore the brain and overcome the problem of limited direct access to the living organ was to draw inferences from observable reactions. One of the earliest inferences, often credited to Paul Broca, is the brain's asymmetry and its functional specialization of the right and left hemispheres, mirrored in another function of our body – namely handedness (Dronkers, Plaisant, Iba-Zizen, & Cabanis, 2007). This insight led to the conclusion that handedness may serve as an indirect marker of cerebral lateralization for language, with right-handers having a dominant left hemisphere for language, and vice versa. A rich discourse discussing the relationship between handedness and language has since then arisen, with many questions remaining unanswered until today: Does hemispheric dominance for language have any conceivable advantages/disadvantages? How has the universal preference for the right hand evolved? How can hemispheric dominance for language or handedness inform us about an individual's likelihood of success in language learning or their general mental ability?

1.1 The Relationship Between Handedness and Language

Handedness can be described as "the individual's preference to use one hand predominantly for unimanual tasks and/or the ability to perform these tasks more efficiently with one hand" (Corey, Hurley, & Foundas, 2001, p. 145). The exact ratio of right- to left-handedness in the human population varies in the literature, but there is a consensus among scientists on approximately 10% of left-handedness in humans (Papadatou-Pastou, 2011). A review by Papadatou-Pastou (2011) found 7.52% left-handers in a sample population of 1.8 million participants and 17.42% non-right handers including left-handers and rare phenomena such as ambidexterity and mixed-handedness. Similar uncertainty dominates the question of the phylogenetic and ontogenetic sources of hand preference. Evidence visible in tools, cave paintings, and preserved bones suggests that handedness and the predominant preference for the right side were already present in Neanderthals (Papadatou-Pastou, 2011). This makes handedness a characteristic that developed very early in our genealogical tree. Concerning the ontogenetic development, Hepper, Shahidullah, and White (1991) were able to prove that foetuses already displayed a clear preference for sucking their right over their left thumb, a trend for handedness that can be observed in human infants at various ages (Goodwin & Michel, 1981).

For long, it has been speculated that handedness must be genetically determined. While it was previously assumed that one single gene codes for hand preference, more recent evidence suggests that a variety of different genes complement each other lead to a certain hand preference. Yet, the exact pattern of inheritance remains largely unexplained, with the only remaining, reasonable inference concluding that the likelihood of having a left-handed baby increases with a left-handed parent (Armour, Davison, & McManus, 2014).

Therefore, the biological basis for the strong right-hand preference in humans has yet to be fully uncovered. What can be said with a reasonable degree of certainty is that the motor function's cerebral lateralization can be found regardless of the preferred handedness. Using fMRI to test motor imagery, Willems, Toni, Hagoort, and Casasanto, (2009) confirmed that performing tasks with the preferred hand activates the cross-lateralized hemisphere, meaning that handedness is left-lateralized in right-handers and right-lateralized in left-handers.

Cross-lateralization also characterises another key process in human anatomy and cognition: language. Connected by the corpus callosum, the right and the left hemisphere of the human brain are each specialized for computing certain cognitive tasks and for processing different sets of information (Papadatou-Pastou, 2011). Language areas are typically found in the left hemisphere, whereas the right hemisphere is said to process visual, spatial, and emotional stimuli. Speculations on atypical right-hemispheric language representation in left-handers have led to antiquated prejudices about left-handedness, with arguably harmful consequences (Corballis, 2014). Considered to be a deviation from the *norm*, left-handedness has often been regarded as a negative trait that is to be corrected. This mind-set gave rise to numerous myths, as well as the stigmatisation of left-handedness still present in society today. Evidence for or against these prejudices is controversial, with recent studies suggesting that disabilities such as stuttering, mental health difficulties, or schizophrenia can be linked to ambidexterity or mixed handedness (Scerri, Brandler, Paracchini, Moris, & Ring, 2011). Yet, the associations between left-handedness and language impairments remain largely unsupported.

1.2 Previous Research on the Relationship Between Handedness and Language

The historic discussion on the idea that language abilities can be assigned to specific brain regions arose in the nineteenth century with Paul Broca's observations about patients suffering from severe speech disorders. His findings suggested that his patients' language impairments were caused by lesions to the left hemisphere, specifically to the region of the lateral frontal lobe (Dronkers et al., 2007). Broca's observations resulted in a first proposed association between brain lateralization and handedness. As all of Broca's patients were right-handed, subsequent predictions on speech representation for left-handers suggested the opposite. This speculation came to be known as Broca's rule (Knecht et al., 2000) and further led to the establishment of the classical model for studying speech representation that relies on data from aphasia patients and is still in common use today (Szaflarski, Holland, Schmithorst, & Byars, 2006).

Since the advent of modern cognitive sciences, a large body of work has been dedicated to the exploration of the relationship between the cross-lateralized functions of language and handedness and the neural organization of language in the human brain. In the 1960s, the Wada test was introduced to explore the functioning of the hemispheres separately and it soon became the standard used in the field. Through the injection of a drug, usually sodium amobarbital, into the patient's (left or right) internal carotid artery, the functions of one hemisphere are suspended (Beimer, Buchtel, & Glynn, 2015). With a complication rate of up to 11%, though, the Wada test remains subject to controversy. Thus, a number of studies have been conducted to research the test's scientific value, as well as to introduce safer, less invasive alternative methods (Binder, 2011; Beimer et al., 2015).

The idea of handedness as a natural marker of cerebral dominance for language is largely discredited among contemporary scientists and has proven insufficient for the clinical setting (Whitehouse & Bishop, 2009). Groen, Whitehouse, Badcock, and Bishop (2013) examined the extent of the connectivity of language lateralization and handedness in the brain. Speaking against Broca's rule, the results indicate that the two functions are largely unconnected in the brain, suggesting that handedness cannot be treated as a reliable marker of hemispheric lateralization for language.

Using fMRI, Pujol, Deus, Losilla, and Capdevilla, (1999) aimed to explore the specifics of the division between right and left hemispheric participation in language tasks. In contrast to previous conclusions the results showed that activation is not as clearly distributed as assumed but rather a gradual process. While atypical activity in the right hemisphere could be established for 24% of left- handed subjects, 14% of them also exhibited left-hemispheric activation and only one subject overall displayed exclusive right-hemispheric lateralization. Knecht et al. (2000) found similar results measuring lateralization in 326 healthy adults with functional transcranial Doppler sonography. The results confirmed a gradual increase of right-hemispheric speech representation in left-handers that was found to be linear with the degree of the participant's left-handedness as assessed by the Edinburgh Inventory. The results were summarized in the formula: "likelihood of right-hemisphere language dominance (%) = 15%-handedness (%)/10" (Knecht et al., 2000, pp. 2512–2518).

A study by Szaflarski et al. (2006) on handedness and lateralization with regard to age also identified language lateralization to the dominant hemisphere as a gradual process. Examining 170 healthy right-handed individuals ranging from 5 to 70 years, they found that while increasing in children and teenagers, there is a gradual decrease of language lateralization to the dominant hemisphere in individuals between 25 and 70 years.

Golestani, Paus, and Zatorre (2002) approached the question of individual differences caused by neural mechanisms from a slightly different angle by assuming that individual language competence levels might be due to differences in brain structure as opposed to brain function or behaviour. The results of their study using voxel-based morphometry supported this assumption by demonstrating variation in brain structure between successful and less successful learners. In her review on the state-of-the-arts imaging research on the relationship between brain structure and language abilities, Golestani (2012) further drew attention to the fact that while studies on brain function could provide insight into the regions of neural activity in response to a certain task, exploring brain structure generated a different set of information as structure was a more stable and less malleable variable. Reviewing studies by Knecht et al., (2000), Dorsaint-Pierre et al., (2006), Propper et al., (2010) and other researchers, Golestani (2012) concluded that right-handed individuals and individuals with a known left-hemispheric dominance for language exhibited a greater left over right structural asymmetry of auditory regions.

Golestani (2012) further addressed the hemispheric specialization for verbal or non-verbal processing. A comparison of studies showed that verbal processing was largely located in the left hemisphere, whereas the non-verbal processing occurred mainly in the right hemisphere. Whitehouse and Bishop (2009), speaking of a complementary specialisation, confirmed this division of information processing for three quarters of their sample population. For the remaining quarter, both functions of verbal and visuospatial processing seemed to be located in the same hemisphere. Despite a number of studies addressing the origins of the brain's labour division, the reason behind hemispheric specialisation remains unclear.

1.3 Cognitive Learning Strategies and Non-verbal IQ

Additional factors to consider in the discussion on cognitive mechanisms underlying individual differences in second language acquisition are cognitive learning styles and non-verbal IQ. There are several models that aim to describe the relationship between language aptitude and cognitive functions. The traditional view by Carroll and Stanley (1959, 1981) treats foreign language aptitude as an independent and separate area that comprises four components: phonetic coding ability, grammatical memory, associative memory, and inductive learning. Contrastingly, Skehan's (1998) *Processing Stage Model* relates aptitude to different levels of cognitive computing of information. According to this concept, general input processing has an influence on phonetic coding ability, and a learner's fluency and output is related to their ability to memorize and retrieve linguistic context. Skehan (1998) further expanded Carroll's four-part model by adding the components of attentional control and working memory. The four stages of SLA processing were defined as (1) noticing, (2) patterning, (3) controlling, and (4) lexicalising.

Another influential model, the Aptitude Complex Hypothesis introduced by Robinson (2007), proposed that there is not one, but many kinds of language aptitudes. Depending on the environmental settings of the language learning process and their diverging conditions, specific learner types may achieve better or worse results. This suggested that the ability to acquire a new language may greatly depend on the compatibility of the learning environment with a person's cognitive patterns. Robinson therefore concluded that these circumstances need to be taken into account when measuring language aptitude.

Contemporary views such as those summarized above stress the dynamic nature of aptitude, defining language aptitude as a trainable and environment-dependent asset rather than a stoic and invariable trait (Biedroń & Pawlak, 2016). In order to capture this dynamic character, Sternberg (2002) introduced an updated testing form, referred to as *Dynamic Testing*, aiming at training participants' language abilities by providing immediate feedback throughout the testing process. This approach, shared by other researches such as Dweck (2006) and Mercer (2012), generated a new series of aptitude tests called the Cognitive Ability for Novelty in Acquisition of Language (CANAL-FT).

Taking a closer look at cognitive styles themselves, the rough categories of field-independent and field-dependent, originally proposed by Witkin (1962) and summarized by Dörnyei and Skehan (2003), provide a starting point in identifying different approaches to information processing. According to the categorization, field-independent learners tend to prefer to study by themselves and break information down into smaller parts, while field-dependent learners usually enjoy social learning settings and adopt a holistic approach when analysing information. The model received vigorous critique, among others from Skehan (1998) who rejected the bipolar view of cognitive styles and its implication that one style may dominate the other. Instead, he proposed that a learner's cognitive ability and cognitive style are two independent qualities. Thus, it is possible that a learner with strong analytical abilities exhibits weak memorizing skills, and vice versa (Dörnyei & Skehan, 2003).

Equally relevant to the discussion on cognitive styles are the questions of whether language aptitude correlates with a person's general IQ and whether aptitude and IQ can be considered unified or rather separate qualities. The perspective that general non-verbal IQ and language aptitude do not correlate with each other is supported by the case of an individual who excelled in acquiring a high level of competence in multiple foreign languages despite a low IQ due to a hydrocephalic brain injury (Reiterer, 2009). A study by Xiang, Dediu, Roberts, Norris, and Hagoort (2012) on the structural connectivity underpinning language aptitude and factors such as IQ indicated that some subcomponents of language correlate more significantly with general IQ than others. Using among others the LLAMA aptitude test series and the Raven's Progressive Matrices, they found positive correlations between IQ and the abilities of vocabulary learning, establishing sound-symbol correspondences, and spatial working memory. Other components of language aptitude, such as sound recognition and grammatical inferencing, seemed less related to general IQ (Xiang et al., 2012).

1.4 Research Questions

Tying language aptitude and the brain together, the current study examines individual differences in the neural mechanisms that underpin our language abilities, using a variety of verbal and non-verbal tests. Specifically, this study investigates the relations between language aptitude and handedness, cerebral lateralization for language, cognitive learning styles, general non-verbal IQ, and game preference.

This study aims to answer the following questions: (1) Does handedness have a statistically visible effect on the language performances of the tested individuals? (2) Does cerebral dominance have a statistically visible effect on the language performances of the tested individuals, and does handedness serve as a marker of cerebral lateralization for language within the tested individuals? (3) Can preferences of learning styles be linked to cerebral dominance and are certain preferences linked to better performances? (4) Is there a correlation between participants' verbal and non-verbal IQ? (5) Are the participants' preferences for either word-based or number-based games reflected in their scores on the respective verbal or non-verbal tests?

2 Methodology

2.1 Participants

The participants were 32 German native speakers (15 male, 17 female) currently living in Vienna, Austria, and recruited from different departments of the University of Vienna. The mean age was 25 years (M = 24.59, SD = 3.1 years), ranging from 19 to 32. Three additional participants were excluded from the study due to either insufficient data provision or a history of corrected handedness.

In order to explore differences between left- and right-handers, participants were chosen based on their handedness. A further factor was background/field of studies, with the goal to arrive at a multifaceted sample pool and increase the diversity of cognitive strategies deployed by the participants. The final sample pool consisted of 16 right-handers (10 females, 6 males) and 16 left-handers (7 females, 9 males) coming from three different backgrounds: arts (12 participants), languages (10 participants), and natural sciences (10 participants). The participants of the arts group (seven females, five males) and studied different degrees ranging from music to art. The participants with language training (six females, four males) all studied English, German or a Romance language. The participants with a background in natural sciences (four females, six males) were either students at the Department of Mathematics or the Department of Physics. To insure the relative validity of the three background groups, one additional female and one right-handed participant had to be excluded from the study due to their double major in English and Mathematics.

All participants had had an average of at least 8 years of English training and at least 4 years of training in a third foreign language (e.g., French, Italian, Spanish, Hungarian, or Czech) in a school setting. All but two participants were currently enrolled post-secondary students: 18 bachelor's degree candidates, 11 master's degree candidates, 1 doctoral degree candidate. The remaining two participants had completed their secondary education and extensive vocational training. All subjects gave their written consent to participate in the study.

2.2 *Instruments*

The test battery consisted of tests and questionnaires measuring the participants' verbal and non-verbal aptitude and exploring their learning strategies and cognitive styles. The study used the following materials: the questionnaire booklet, two verbal-tests, and one non-verbal test.

2.2.1 Questionnaires

Two complementary questionnaires collected information on the participants' linguistic background, their learning strategies, and their presumable hemispheric dominance. The obtained data provided the foundation for a qualitative analysis.

The first questionnaire, Basic_Quest, collected basic information on the participants. It enquired about gender, age, handedness, educational background, training experiences, and current employment. In order to establish a linguistic profile, participants were asked to give a detailed account of their language background and rate their competences using a scale from 1 (very low competence) to 10 (native speaker). Additionally, participants indicated whether language learning had occurred in an institutional setting, a domestic setting, through contact with native speakers, or as a result of residency in a foreign country (>2 months). Several open questions addressed the participants' personal motivation for learning foreign languages and enquired about their usual study habits, as well as their personal preferences concerning tasks or games that deal with either verbal or non-verbal content. The participants' opinions on their abilities were recorded via self-assessment questions concerning their creativity, their ability to think logically, their preference for either languages or natural sciences in school, and the time devoted each week to creative activities. These questions used a 7-point scale. The questionnaire was refined by adding two questions on the participants' handedness: first, whether participants had left-handed relatives, and second, whether their handedness had been corrected in any way. The final version of the Basic_Quest contained 20 items, 8 addressing the participants' personal and linguistic background and 12 establishing a character profile.

The second questionnaire, Hem_Quest, was an adapted version of an online questionnaire seeking to determine the participants' dominant hemisphere. Traditionally, invasive and costly tests such as the Wada test are used to establish hemispheric dominance in clinical patients. Alternative and less invasive measurements have been proposed and tested but a non-invasive, reliable, and publicly accessible method is still far in the future. In popular culture, handedness therefore still serves as a frequently used measurement of hemispheric dominance and countless websites invite people to assess their cerebral dominance for language by answering lengthy sets of questions. This study used a short collection of ten single-choice questions to gather information on the participants' thinking strategies, cognitive patterns, and character traits. While the answers could not provide a scientific estimation of hemispheric dominance for language, they allowed for an interpretation of the subjects' probable tendency towards a certain cognitive style and complemented the results of the Basic_Quest.

Both the Basic_Quest and Hem_Quest were printed and completed with a front page stating the nature of the project, its extent, and included a consent form. Each booklet of questionnaires was marked with a number to ensure the participants' anonymity.

2.2.2 Language Aptitude Testing

This study used a combination of the MLAT V and the LLAMA F to test the participants' aptitude for vocabulary learning and grammatical inferencing. The MLAT, short for Modern Language Aptitude Test, was developed by Carroll and Sapon in 1959 to compensate for the insufficiency of general intelligence tests to predict a person's likelihood of success in acquiring a new language (Sparks & Ganschow, 2001). The MLAT series has since become the most influential language aptitude test and has acted as a basis for subsequent batteries and research (Biedroń & Pawlak, 2016). The design of the original test is fourfold and assesses the subcomponents of phonetic coding ability, grammatical sensitivity, inductive language learning ability, and associative memory (Dörnyei & Skehan, 2003).

This study used the test component MLAT V, which measures a participants' ability to form connections between vocabulary of a familiar and a foreign language. Participants had to memorize 24 Kurdish words in 2 min and reproduce them afterwards by choosing between four options for the English equivalent. A possible drawback of using the MLAT V for a subject pool of native-German speakers was that the participants needed to link words of a foreign language to words of their L2 instead of their L1. However, the test asked for simple everyday words all participants could be expected to be familiar with.

A very successful alternative to the MLAT series is the LLAMA series, developed by Paul Meara (2005) from the University of Swansea. In the official LLAMA Manual, Meara (2005) states that the test design, which was influenced by Carroll's work, aims to provide an alternative for non-English native speakers and therefore works with pictures and a constructed language. After extensive revision work, the current test series consists of four individual tests that can be downloaded free of charge from the internet and completed on a home computer (requirement: Windows). The LLAMA vocabulary task, which correlates with the MLAT V paired association task, solves the inherent L1-related issues by relying on pictures and made-up words.

The second verbal test used in this study was the LLAMA F, which assesses the ability of grammatical inferencing. During this test participants need to understand the grammatical system of a language and are given 20 pictorial items showing a variety of simple symbols. The items are accompanied by a short description in a constructed language. The challenge is to infer certain grammatical patterns and rules of the new language by comparing similarities and differences between the pictorial content and the descriptions. Afterwards, participants had to link a picture with the correct description by deciding between two alternatives. The focus of the LLAMA F is mainly on word order and agreement features. The LLAMA Manual also provides a reference scale to interpret the scores. Results between 20% and 45% are considered an average score, results of 75% and higher an outstanding performance only achieved by few participants.

2.2.3 Non-verbal IQ

The participants' non-verbal IQ was measured with the Raven's Progressive Matrices, one of the most widely used measures of cognitive ability. The test, initially developed by Spearman in 1927 and revised and published by John Raven in between 1938 and 1941, relies on visual stimuli and dispenses with linguistic input altogether. The Raven's Progressive Matrices was constructed to assess a person's general cognitive ability, which represents the ease with which a person can answer questions and infer answers from given material, and their reproductive ability (Raven, 2000). Several standardisations based on data obtained from studies carried out in Great Britain (Adams, 1952; Raven, 1938 (published in 1941)) and Germany (Kratzmeier & Horn, 1980) showed a relative stability of the test results and have generated regularly updated norms, making the Raven's Progressive Matrices a well-validated tool to assess a person's general mental ability (Raven, 2000). The test is therefore a popular choice in a variety of settings, from educational institutes to clinical centres.

This study used a version of the Raven's Progressive Matrices consisting of 48 black-and-white items that become increasingly demanding. The subject was presented with one pictorial item at a time and was asked to complete the pattern's missing piece by choosing from a set of multiple-choice items.

2.3 Procedure

All subjects were given the same instructions and were tested individually. In order to keep the participants motivated and engaged, the test battery was designed to begin with the shortest and end with the most time-consuming task. The first of three test blocks consisted of the Basic_Quest and Hem_Quest. The second block contained the two verbal tests, the MLAT V and LLAMA F. The 2 min of the study phase were digitally timed and generally there was no time limit for the testing phase. The LLAMA F consisted of a study phase during which participants were given the option to take notes, and a testing phase which was completed on a computer. The third and final block consisted of Raven's Progressive Matrices. The graphical items were shown to participants on a computer screen, and the answers were filled in manually on a test sheet. There was no time limit, and subjects had full control over the computer and the time spent on individual items. The scores for all tests were determined by adding up the number of correct answers given. Errors or omissions did not result in the loss of points. Throughout the testing process, remarks by participants regarding their own learning styles inspired an additional question concerning cognitive strategies. Immediately after completing MLAT V, ten participants were asked to reflect and briefly describe their approach to completing the memorization task.

3 Results

The results were analysed statistically using SSPS 22. The alpha level for all statistical tests was $p < .05$. As a first step, descriptive statistics for the total sample were computed. The questions with qualitative value (Basic_Quest, Hem_Quest) were analysed separately from the quantitative test results.

3.1 Questionnaires

3.1.1 Results of the Basic_Quest

The first step in answering this paper's research questions was to establish the participants' linguistic profile, which allowed a subdivision of the participants into smaller groups. The Basic_Quest provided a vast amount of information, not all of it relevant to the current study. The following report summarizes these results and includes additional information that was captured in order to gain a full picture of the subject pool and to rule out other factors, such as performance differences according to gender.

The results of the Basic_Quest showed that the participants whose declared area of study was language generally spoke more languages (M = 3.4, SD = 0.84) than participants studying natural sciences (M = 3.0, SD = 1.0) or arts (M = 2.7, SD = 0.89). The difference, however, was not significant. The numbers of languages spoken by the participants ranged from 2 to 5, with 10 participants speaking 2 languages, 13 participants speaking 3 languages, 6 participants speaking 4 languages, and 3 participants speaking 5 languages. The most common second language of the sample group was English, with every participant having had between 8 and 15 years (M = 10.19, SD = 2.3) of training at an educational institute (school, university, summer schools). Spanish and French shared the place of the most commonly spoken L3, with 75% of participants having learned one (or both) of the languages in a school setting. Other languages spoken by the participants included Hungarian, Czech, Icelandic, Swedish, Arabic, Korean, Polish, and Turkish. None of the participants spoke Kurdish.

Over half of the participants (54.5%) who had acquired a foreign language in their spare time and outside of an educational setting indicated that their motivation was driven by the desire to be able to better communicate with native speakers of the respective languages. Other answers included the wish to learn the partners' native language, curiosity, and a general interest in the customs, language, and traditions of other language communities.

Correlations between the number of languages spoken and the test scores for the MLAT V, LLAMA F, and Raven's Progressive Matrices (henceforth RPM) revealed that language proficiency and the motivation to acquire languages outside an educational setting did not influence language competence as measured by the individual tests (MLAT V: $r = -.053$, $n = 32$, $p = .774$.; LLAMA F: $r = -.02$, $n = 32$, $p = .912$; RPM: $r = -.277$, $n = 32$, $p = .124$).

The distribution of handedness for the final sample showed a slight difference concerning gender, as 58.8% of females compared with only 40% of males were right-handed. This indicated that the majority of male subjects were left-handed. The follow-up question enquiring about the tendency of handedness in the participants' close family revealed that 75% of left-handed subjects had left-handed relatives, whereas this was only true for 12.5% of right-handed subjects. A Pearson product-moment correlation for the data confirmed a highly positive correlation between the variables of left-handedness and left-handed relatives, $r = -.630=$, $n = 32$, $p = .000$.

In order to explore the research question (3) on the correlation between learning styles and handedness/cerebral dominance/performance, participants were divided into two groups of different learning styles, the creative style group and the organized style group. The division was based on the participants' answers indicated in the Basic_Quest. Participants associated with the creative style stated a preference for social learning environments, approaching tasks by first looking at the whole picture and only later at the individual details, and using visual imagery and creative strategies such as drawing pictures and sketches or creating a story out of important elements. Participants associated with the organized style reported a preference for studying alone, approaching tasks in an organized manner and using strategies such as making lists or looking at the individual elements in detail before considering the whole. Participants who combined elements from both styles were categorized by considering the self-assessment of their creativity/logic.

According to the results, a majority of the participants (56.25%) used an organized style. An independent sample t-test showed that the learning styles had no significant effect on the participants' performances on the tests (MLAT V: $t(30) = -.621$, $p = .539$; LLAMA F: $t(30) = -1.527$, $p = .137$; RPM $t(30) = -.531$, $p = .619$).

The creative group consisted of 12 participants with right-hemispheric dominance as assessed by the Hem_Quest, and 2 participants with left- and balanced hemispheric dominance, respectively. The organized group consisted of 5 participants with right-hemispheric dominance as assessed by Hem_Quest, 11 participants with left-hemispheric dominance, and 2 participants with balanced dominance. Concerning handedness, half of the right-handed participants used the creative style and the other half the organized style, whereas a majority of ten left-handers associated themselves with the organized style and only six left-handers with the creative style.

Relevant to research question (5) on game preferences, a majority of 53.2% of participants preferred non-verbal over verbal games in a situation where they could choose between a Sudoku puzzle (non-verbal) and a crossword puzzle (verbal). Concerning gender and game preference, 74.7% of females and 40% of males opted for Sudoku, indicating a strong preference for the non-verbal task among female participants. A comparison between the three student groups showed that an equal amount of language students and natural sciences students opted for the Sudoku (60% respectively), while a majority of the artists chose the Crossword puzzle (58%).

While the games were equally popular among left-handed participants, 56.25% of right-handed participants settled on solving a Sudoku. Participants with a right-hemispheric dominance as assessed by the Hem_Quest preferred solving a crossword (58.8%) to Sudokus (41.2%), whereas a majority of participants with left-hemispheric dominance chose a Sudoku (66.67%). From the three participants with balanced hemispheric activity, two chose Sudokus and one chose a crossword.

A t-test was computed to assess the relationship between game choice and the test results. The effect of game choice is different between the three tests. Participants who preferred the Sudoku scored significantly higher on the LLAMA F ($M = 64.7$, $SD = 19.08$) than participants who preferred the crossword puzzle ($M = 46.0$, $SD = 15.95$), $t(30) = -.992$, $N = 32$, $p = .006$. These results suggest that practice in non-verbal games provided a significant advantage to the participants in completing the grammatical inferencing task. Further, there is a visible though not statistically significant trend for participants favouring the Sudoku to also score higher on the MLAT V, $t(30) = -1.814$, $N = 32$, $p = .08$. The game choice did not seem to have any influence on how well participants scored on the non-verbal RAVEN, $t(30) = -.992$, $N = 32$, $p = .329$.

3.1.2 Results of the HEM_Quest

The results of the Hem_Quest provided the basis information needed to explore the research question (2) on the effect of cerebral dominance on language performance and whether cerebral dominance might be linked to handedness.

According to the questionnaire, a majority of 53.1% of the participants displayed a dominant right hemisphere. Among the remaining participants, 37.5% displayed a dominant left hemisphere and 9.4% were balanced. Concerning handedness, the ratio of right- and left hemispheric dominance for right-handers was 69–25% (right: left) with one subject (6%) tested as balanced. Left-handers showed results of 37.5–50% (right: left) and two subjects (12.5%) tested as balanced. An analysis of brain dominance between the three groups of art students, language students, and natural sciences students showed that the majority of both the art students and language students indicated a dominant right hemisphere with 58% and 70% respectively, whereas the left hemisphere was predominant for participants studying natural sciences (70% left and 30% right).

While an equal amount of male participants showed a right or left hemispheric dominance (46.6% respectively, and 6.6% balanced hemispheric activity), the majority (58.8%) of the female sample group displayed right hemispheric dominance. Among the remaining female participants, 29.4% indicated a left hemispheric dominance and 11.8% showed balanced cerebral activity.

3.2 *Verbal and Non-verbal Tests*

The results of the verbal and non-verbal tests established the participants' language performance and were crucial to answer the research question (4) on whether a correlation between verbal and non-verbal IQ can be established.

As a first step, the scores for all three tests were calculated. For the entire sample, the scores for MLAT V ranged from 5 to 24 (of 24 possible points), with $M = 17.19$, $SD = 6.03$. For LLAMA F, participants achieved between 20% and 100%, with $M = 55.93$, $SD = 19.81$. The maximum score achieved for the RPM was 46 (out of 48 possible points), with a minimum score of 5 points. The mean score was $M = 37.18$, $SD = 8.14$.

In regard the research question (1) on whether handedness influences language performance, the group of right-handers achieved higher results for both verbal tests (MLAT V $M = 18.5$, $SD = 6.7$; LLAMA F $M = 58.8$, $SD = 22.7$) and for the non-verbal test (RPM $M = 38.1$, $SD = 5.3$) than the group of left-handers (MLAT V $M = 16$, $SD = 5.1$; LLAMA F $M = 53.1$, $SD = 16.6$; RPM $M = 36.3$, $SD = 10.3$). An independent sample t-test showed that these differences, however, were not significant.

Concerning the research question (2) on the effect of hemispheric dominance as assessed by the Hem_Quest on the test scores, a one-way between subjects ANOVA was computed. Multiple comparisons between the groups showed that there was no advantage for any group of individuals according to their hemispheric dominance.

A Pearson product-moment correlation coefficient was computed to assess the relationship between the verbal tests and the RPM and answer the research question (4) on verbal and non-verbal IQ. There was a positive correlation between the MLAT V and the RPM, $r = .494$, $N = 32$, $p = .004$. The LLAMA F and RPM correlated with similar high significance, $r = .383$, $N = 32$, $p = .031$. Among themselves, the verbal tests correlated strongly with $r = .578$, $N = 32$, $p = .001$.

To further look into the matter of verbal and non-verbal IQ and language performance, a one-way between subjects ANOVA was conducted to compare the effect of educational background and career choice on the test scores between the groups of art students, language students, and natural science students. There were no significant effects on the overall test scores at the $p < .05$ level. However, the variable of educational background and career choice had a strong tendency towards statistical significance on the scores for the MLAT V, [$F(2,29) = 3.226$, $p = .054$], and a weaker tendency on the scores for the LLAMA F, [$F(2,29) = 980.104$, $p = .078$]. Because of the tendencies towards statistical significance in this sample, a post hoc test was computed to gain a better understanding of the effect of educational background and career choice on the individual test results. Post hoc comparisons using the Tukey HSD test indicated slight differences in performance between the groups of art students and language students. These differences were not significant at the $p = <.05$ level. Taken together, these results suggest that educational background and career choice had no measurable effect on the participants' performances for the MLAT V, LLAMA F, and RPM.

3.3 Additional Findings: Qualitative Exploratory Analysis

Throughout the process of data collection, ten participants were encouraged to reflect on their learning strategies and to briefly describe their approach to completing the task. Immediately after finishing the MLAT_V, they were given a short reflection time so that they could write down their thoughts in the questionnaire booklet. The answers given by the participants are summarized in Table 1. The table also indicates the following factors: whether the answer was given multiple times, the participants' handedness (R = right, L = left), hemispheric dominance as assessed by the Hem_Quest (hem_dom: R =right, L = left, B = balanced), and the achieved results on the MLAT_V (scores).

As the participants were chosen randomly, the ratio of right- to left-handers and right- to left-hemispheric dominance was incidental. On the whole, it can be said that there are clear differences in the answers given by participants with a right- or left-hemispheric dominance, indicating a correlation between hemispheric dominance as assessed by the Hem-Quest questionnaire and approaches of creative or organized manner.

Table 1 Individual learning strategies for MLAT_V

Q: How did you memorize these words?	Handedness	hem_dom	Score (0–24)
Made up a song with the words (2×)	R, L	R, R	16, 21
Came up with a story	R	R	20
Mnemonics	R	R	21
Read through list once very slowly	R	L	24
Read through list often in different order (3×)	R, L, R	L, L, L	17
Self quizzing (2×)	L, L	L, L	8, 19

4 Discussion

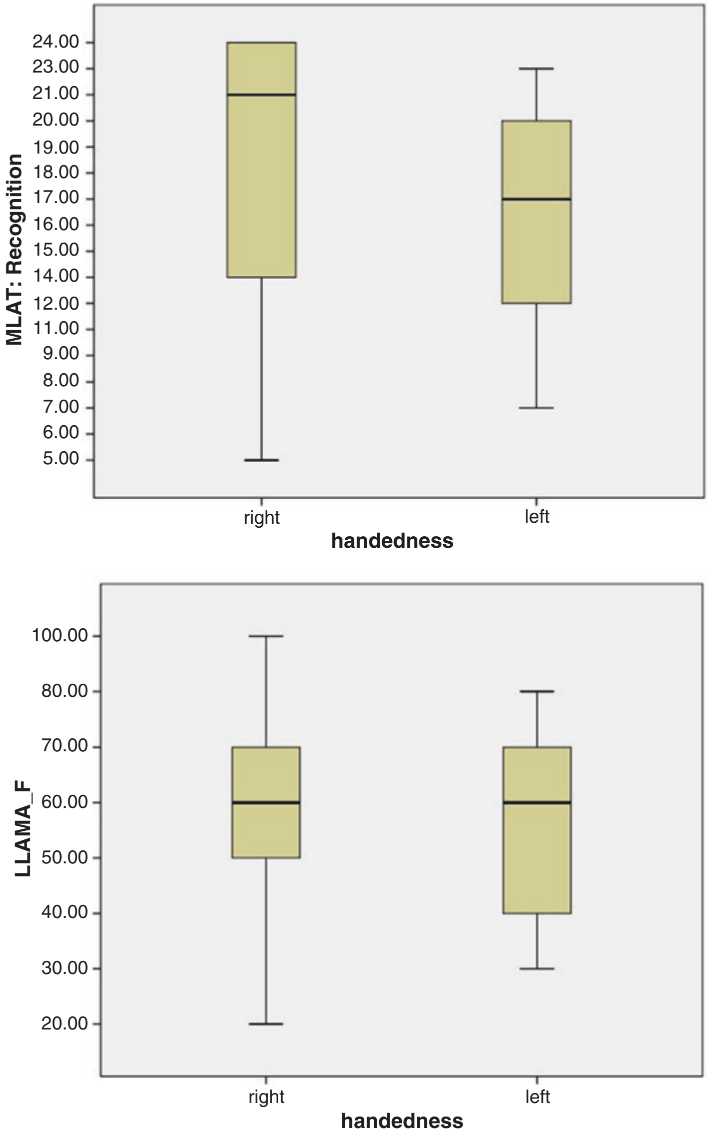
The current study examined individual differences in the neural mechanisms that underpin our language abilities. Specifically, this study aims to answer the following questions: (1) Does handedness have a statistically visible effect on the language performances of the tested individuals? (2) Does cerebral dominance have a statistically visible effect on the language performances of the tested individuals, and does handedness serve as a marker of cerebral lateralization for language within the tested individuals? (3) Can preferences of learning styles be linked to cerebral dominance and are certain preferences linked to better performances? (4) Is there a correlation between participants' verbal and non-verbal IQ? (5) Are the participants' preferences for either word-based or number-based games reflected in their scores on the respective verbal or non-verbal tests?

In this section, the aforementioned results will be discussed in greater detail and explored in relation to the theoretical framework, with the aim to answer this paper's research questions.

4.1 The Effect of Handedness on Language Performance

Answering research question (1) on handedness and language performance, the results of the current study suggests that there is no significant effect of handedness on language competence. As expected, neither the group of right-handed nor left-handed participants demonstrated decisive superiority over the other in regard to their language performances. However, an interesting point to mention is the variation in the scoring spectrum between the two groups. As illustrated in Figs. 1 and 2, right-handed participants claimed the highest and the lowest scores for both the MLAT V and the LLAMA F. This gives the group's results a broad range, from a minimum of 6 to a maximum of 24 out of 24 points in the case of the MLAT V, and from 20% to 100% with regard to LLAMA F.

In contrast, the group of left-handers produced more consistent results with a lower deviation from the mean for both verbal tests. The group of left-handers therefore exhibited more stability and internal consistency in their results, whereas the scoring spectrum of the right-handers was characterized by extremes. This trend is



Figs. 1 and 2 Boxplots illustrating the scoring spectrum for the MLAT_V and LLAMA_F for the groups of right- and left-handers

partially continued in the results for the non-verbal task, with right-handers claiming the highest scores of 44 and 46 out of 48 points (Fig. 3). However, a left-handed participant scoring the lowest result for Raven's Progressive Matrices with only 5 out of 48 points breaks the pattern observed in the verbal tests. The huge discrepancy between the individual's and the remaining participants' performances qualify the left-handed subject as an outlier, together with another left-handed participant

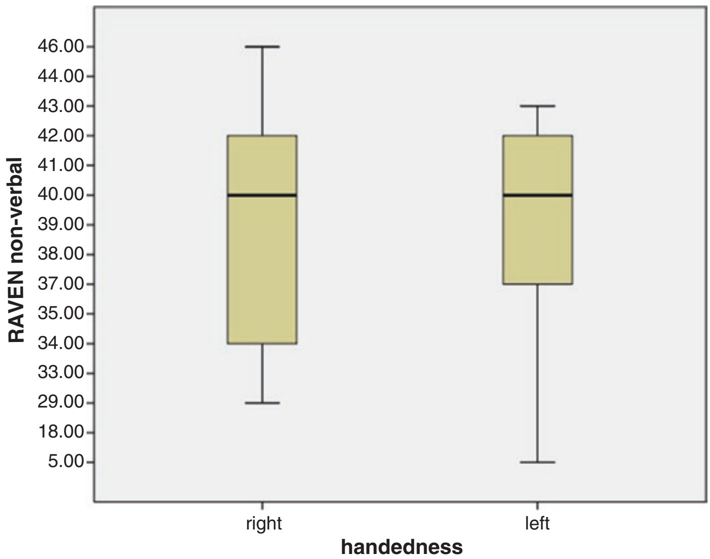


Fig. 3 Boxplot illustrating scoring spectrum for RAVEN for the groups of right- and left-handers

who achieved the second lowest result. A closer analysis of the participant in question (left-handed, male, language background, right hemisphere) shows that he scored at the lower end for the MLAT_V (8 out of 24, fifth lowest overall performance for MLAT_V), but within the average for the LLAMA_F. Due to the average results on the RAVEN his data was not excluded from the analysis.

To summarize, the results indicated variation in the scoring spectrums of the groups of left- and right-handers concerning the verbal tests, yet hand preference presented no overall significant advantage or disadvantage on the participants' language performances. This conclusion is largely consistent with predictions from previous research showing that left-handedness or atypical language representation in the right hemisphere do not present inhibition for individuals or require manipulation and treatment (Corballis, 2014).

Findings like these are of principal importance to the discourse on handedness, as they suggest dismantling prejudices against left-handedness, a trait that despite its rare occurrence still affects 10% of the human population. Compared to the amount of research conducted on right-handedness, the phenomenon of left-handedness is largely unexplored and has therefore been associated mostly with disadvantages (Johnston, Nicholls, Shah, & Shields, 2009). The stigmatization of left-handedness is further fuelled by controversial results. For example, Johnston et al., (2009) conducted a large-scale and longitudinal study on handedness in Australian children, which indicated that right-handed children of the sample group had a clear cognitive advantage over left-handed and ambidextrous children. Corballis (2014), on the other hand, reviewing the origin of brain asymmetries and their implications, pleads for the irrelevance of handedness for a person's cognitive abilities and names examples of famous left-handed people, such as Leonardo da Vinci and Barack Obama.

It is this lack of quantity and quality in research on left-handedness that has led to the belief that a divergence from right-handedness requires manipulation. Considering the results of the current study regarding the research question (1), there is no evidence that handedness had a measurable influence on the participants' performances. Instead, observed differences are more likely due to other factors, such as personal cognitive styles and learning strategies.

Another interesting factor regarding handedness as explored in this study is the highly significant correlation between left-handed participants and left-handed family members. Research into the origin and development of hand preferences strongly suggests that left-handedness is linked to environmental factors and genetics (Szaflarski et al., 2006). A study by Dehaene-Lambertz, Dehaene, and Hertz-Pannier (2002) showed that language lateralization is measurable already in early childhood at a point at which the infant has been subject to minimal social and cultural influence. This data strongly favours genetics as the essential factor determining hand preference. In accordance with these findings, the results of the current study clearly highlight the genetic link, as 12 out of 16 left-handed participants reported to have left-handed relatives, whereas this is only true for two out of 16 right-handed participants. Follow-up questions conducted after the study revealed that the two right-handed participants with left-handed relatives were sisters, with a third left-handed sister who was also part of the sample pool.

4.2 The Effect of Cerebral Dominance on Language Performance and Cerebral Dominance As a Marker of Handedness

The first step in answering the research question (2) on cerebral dominance was to subdivide the subject pool based on the results of the Hem_Quest. The results of the questionnaire show a slender majority (53.1%) of right-hemispheric dominance in the total sample group, meaning that just over half of the participants identified themselves with activities and coping mechanisms that are connected to the right rather than to the left hemisphere. Additionally, more right-handers tested for a right- than a left-hemispheric dominance, and more left-handers for a left- than a right-hemispheric dominance. These results raise several questions, as they do not agree with previous literature.

Despite uncertainty concerning the exact ratio of right- and left- handedness in the human population, scientists agree that a vast majority of approximately 90% of humans exhibit a preference for using their right hand (Papadatou-Pastou, 2011). Broca's rule, linking handedness to cross-lateralized language areas in the brain, would therefore predict a dominant left hemisphere for language in 90% of humans.

The results of the current study do not align with these predictions. In fact, a majority of the participants tested for an atypical right language representation, suggesting the opposite.

However, it has to be noted that the participants' hemispheric dominance was established by the use of a non-validated questionnaire that will not stand up to serious scientific scrutiny. An abundance of similar questionnaires can be found on the internet, offering a quick analysis and interpretation of cerebral dominance and satisfying the growing demand of casual interest. The questionnaires' popularity presumably arises out of their informative content that not only categorises the participant's character based on the provided input, but further generates information on typical traits, strong points, and weaknesses of a dominant right or left brain, therefore helping interested individuals reflect on their capabilities and techniques.

Thus, evaluation of the Hem_Quest results leads to two main conclusions. First, despite its questionable validity, the findings support the view that handedness is not a reliable marker of cerebral lateralization for language. Even though a number of studies have questioned and confirmed the validity and invalidity of this 'golden standard', a majority of projects published in the first decade of the 2000s still relied on handedness as an indicator for language lateralization (Groen et al., 2013), mainly because of a lack of suitable alternatives to the invasive method of the Wada test. An increasingly large body of research therefore concentrates on the search for non-invasive alternatives, with promising results including the use of fMRI (Binder, 2011) or magnetencehaplography (MEG) or near-infrared spectroscopy (NIRS) among others (Pelletier, Sauerwein, Lepore, Saint-Amour, & Lassonde, 2007).

Secondly, the debatable results of the Hem_Quest disagreeing with predictions based on previous literature confirm that the medium of a questionnaire does not suffice in determining cerebral dominance for language. The Hem_Quest does, however, provide a valuable and deeper insight into the participants' cognitive strategies. The Hem_Quest and its online equivalents should therefore be regarded as a source of information on learning styles, rather than measuring cerebral lateralization for language.

4.3 Learning Strategies in Relation to Handedness, Cerebral Dominance, and Language Performance

The concept of 'different learning styles' in SLA, subject matter of research question (3), describes diverging tendencies of favouring one set of cognitive strategies over others (Dörnyei & Skehan, 2003). This perspective suggests that no style is superior over the others – rather, every strategy has its own advantages and drawbacks. The findings of the current study align with this view, as the variable of learning styles had no observable effect on the participants' performances. The rough categories of 'creative style' and 'organized style' follow Witkin's (1962) view of the field dependent and field independent learner, to the extent that the 'creative style' is associated with a social learning setting and a holistic approach, and the 'organized style' with independent studying and an analytical approach. Concerning the current study, this distinction produced two 'rough' categories, because the distribution relied on a relatively small number of answers that could only capture a fracture of the participants' predisposition towards certain learning styles. Additionally, a dual distinction forced participants into groups on two extreme ends of a spectrum. Therefore, the categories of 'creative style' and 'organized style' are rough distinctions rather than finely tuned categories of learning styles.

A correlation of learning style with the participants' hemispheric dominance as assessed by the Hem_Quest lends some credibility to the categorisation. 70.59% of right-brained participants leaned towards the 'creative style', while an overwhelming majority of 91.67% of left-brained participants leaned towards the 'organized style. This follows the perspective of our brain's 'complementary specialisation' (Whitehouse & Bishop, 2009), with the right hemisphere being the 'creative side' that processes subjective, intuitive and emotional stimuli. By contrast, the left hemisphere and 'logical side' favours the rational, analytical, and objective stimuli.

This distinction is also visible in the answers from the ten participants who reflected on their learning strategies after completing the MLAT_V. As the summary of the answers clearly showed, participants tested for right hemispheric dominance preferred creative learning strategies that involved, for example, music and narrative elements. Left-brained participants, on the other hand, reported using analytical and rational learning strategies.

To answer research question (3), it can therefore be said that the qualitative and quantitative analysis of the participants' learning styles indicates that the participants' learning styles had no significant effect on their language performances, despite the correlations between cerebral dominance and learning strategies connected to either the right or the left hemisphere. Instead, the findings support Dörnyei and Skehan's view (2003) that each learning style can be characterized by both advantages and disadvantages.

4.4 The Correlation of Verbal and Non-verbal IQ

Turning to research questions (4), highly positive correlations between the RAVEN and the two verbal tests indicate a high correspondence between a person's verbal and non-verbal IQ. The results suggest that participants who have an easy time completing patterns and inferring abstract rules from given information are likely to also exhibit good memorization skills. Furthermore, the high correlation between the RAVEN and the LLAMA_F ($r = .383$, $N = 32$, $p = .031$) suggests that there is a connection between the abilities of making sense of abstract or pictorial structures and structures relying on linguistic content.

The positive correlations may in part be due to the similar test designs of the RAVEN and the verbal tests measuring similar cognitive abilities. The RAVEN aims to establish a person's general cognitive capacity ('g'), which comprises the abilities to "make meaning out of confusion, […] to generate high-level, usually nonverbal, schemata which make it easy to handle complexity" and to "absorb, recall, and reproduce information that has been made explicit and communicated from one person to another" (Raven, 2000, p. 2). The LLAMA_F, a grammatical inferencing task, tests for the first component of pattern recognition, with the difference being the focus on verbal instead of non-verbal structures. The MLAT_V can be said to test for the second component, as it assesses the ability to "learn and retain associations between words of a new language and their meaning in English" (Language Learning and Testing Foundation, 2014). The tests, while measuring a person's non-verbal IQ and verbal abilities respectively, therefore draw on similar cognitive resources, which explains the high correlations between the three tests.

The results of this study partially agree with the findings from Xiang et al. (2012) concerning the positive correlations between vocabulary learning and IQ. While Xiang et al. (2012) did not establish a positive connection between grammatical inferencing and IQ, the results of this study suggest otherwise.

Therefore, the current study showed that language abilities and general IQ are connected by a complex and multifaceted relationship. Future research on the exact nature of this relationship will help to explain diverging results and exceptional cases.

4.5 Game Preferences and Language Performance

The last research question (5) addressing the participants' preferences regarding games of either verbal or non-verbal content generated highly interesting results. The participants were presented with two options: a crossword puzzle (verbal) or a Sudoku puzzle (non-verbal). These games were chosen because they demand diverging processing strategies and a different kind of input from their players. The crossword puzzle relies on linguistic input and on a substantial mental lexicon, whereas the Sudoku puzzle requires a holistic approach and good visuospatial skills, meaning the ability to process and complete a large amount of information as well as determine an item's location in space. While the visuospatial aspect also plays a role in solving a crossword, a player's probable success depends considerably less on this ability than for solving a Sudoku.

As a large amount of research established, these sets of abilities are located in different areas of the brain, with language typically lateralized to the left, and visuospatial skills to the right cerebral hemisphere (Golestani, 2012; Whitehouse & Bishop, 2009). The current study therefore aimed to compare the participants' game preferences to their hemispheric dominance as assessed by the Hem_Quest, expecting that the popularity of the crossword puzzle would be higher for left-hemispheric participants, and the Sudoku for right-hemispheric participants. Interestingly, the results suggest the exact opposite, as a majority of participants with right-hemispheric dominance opted for solving the verbal puzzle, and vice versa. Additionally, over half of the language students preferred the Sudoku to the crossword puzzle. Thus, the findings of this study indicate that the participants' game choices were influenced by other factors than their usual cognitive strategies for solving tasks or than their chosen professional environment.

Further, there was a significant difference between genders, as the Sudoku was more popular with female participants and the crossword with male participants. This finding contradicts stereotypical assumptions that link numbers and mathematics to the male, and language to the female domain.

Bivariate testing showed that there were significant differences between the scores achieved by the participants opting for the Sudoku to those opting for the crossword. Against probable expectations, participants preferring the non-verbal game scored higher on both verbal-tests. This effect could not be observed for the non-verbal RAVEN, raising the question why practice in non-verbal games provided participants with an observable advantage in solving verbal tests. As the LLAMA_F focuses on word order and word agreement, it arguably demands high visuospatial skills. The results clearly indicate that the LLAMA_F draws on a very similar skill set as a Sudoku does, in spite of the different content and nature of the task's building blocks. It therefore seems that in order to learn a new language the ability to infer and deduce abstract patterns from given information may be more important than a substantial lexicon or a good word memory.

The main findings concerning research question (5) on game preferences are, first, that the game choice did not correlate with the participant's hemispheric dominance as assessed by the Hem_Quest. Instead the results showed that left-brained participants and language students preferred the non-verbal game alternative. Second, participants opting for the Sudoku scored significantly higher on the LLAMA_F, suggesting that practice in inferencing tasks regardless of their content can provide a measurable advantage in second language learning.

4.6 Potential Limitations of the Study

This study is subject to several possible limitations and weaknesses. First, the Raven's Progressive Matrices is a rather time-consuming test that took the participants between 30 min to over an hour to complete. It was therefore too long for some participants to do without fatigue. Participants were allowed and also encouraged to take short breaks and spend as much time as needed on the individual items. Nevertheless, it was obvious that some subjects lost interest or patience in the careful completion of the task, especially as the items became increasingly difficult to solve. The possibility of guessing can therefore not be ruled out. Additionally, the time of day when the tests were conducted differed between the participants. Some participants tested in the evening reported that the late timing had a negative effect on their attention span. To minimize instances of impatience and inattentiveness, the option to complete the test at home was made available. Thereby, participants were able to take the test at a time when they felt at an intellectual peak and in more than one sitting, reducing the risk of guessing answers. The obvious drawback of this option was the lack of control over the testing situation and a rising dropout rate.

Second, the MLAT_V was developed for native speakers of English, yet all participants of this study acquired English as an L2. The MLAT_V was chosen over its equivalent (LLAMA_B) from the LLAMA test series for two reasons: first, because the MLAT_V is the more timesaving option, and also because the LLAMA_B has to be downloaded and can only be completed on a computer (Windows). Furthermore, the language barrier was not considered to jeopardize the scientific validity of this study, as all participants had a high level of education and an average of 10 years of training in the English language. Additionally, the MLAT_V asks for simple and everyday vocabulary that the participants could be expected to be familiar with.

However, four participants (three males, one female) showed signs of confusion and distress throughout completing the task and achieved very low scores (5, 7, 7, and 8 from 24 points). When asked, the four participants reflected that even though they had years of language training, their contact with the English language had stopped after graduating from high school. All four participants were art students and reported using English only on very few occasions. Therefore, the LLAMA_B, which uses pictorial elements and made-up words, might have been a better choice for this study.

Another weakness of this study is the analysis of handedness and cerebral lateralization for language. Due to time pressure and lack of resources, the participants' degree of handedness was not confirmed by any measurement, such as the Edinburgh Handedness Inventory. To minimize the possibility of balanced handedness, three participants who indicated that their handedness had been corrected in any way were excluded from the study. As cerebral lateralization for language is very hard to determine by means that do not include neuro-imaging techniques, the Hem_Quest was used as an alternative to capture the participant's probable tendency towards specific thinking patterns that can be assumed to be linked to a dominant right or left hemisphere. The results of this non-validated measurement of hemispheric dominance can be expected to diverge significantly from reality. This is why the phrase 'as assessed by the Hem_Quest' is continuously added to the results of hemispheric dominance in the current study.

The question regarding game preference was originally added to the Basic_Quest without high expectations of generating interesting results. However, the findings of this study show measurable differences in the performances of the participants favouring the Sudoku to the crossword. It would therefore have been interesting to establish and reconfirm the participants' game preferences through multiple questions on games of either verbal or non-verbal content and to further distinguish between games that participants prefer and games that participants are best at.

5 Conclusion

To conclude, the current study examined individual differences in language aptitude. Specifically, it explored the relationship between language performance and the variables of handedness, hemispheric dominance, cognitive learning strategies, non-verbal IQ, and game preferences. The findings of this study strongly support the view that handedness has no measurable effect on second language learning success. Likewise, hemispheric dominance for either the right or the left hemisphere as assessed by the non-validated Hem_Quest posed no significant advantages for either group. A qualitative analysis of the participants' learning strategies suggests that different styles are equally legitimate and tend to correlate with the participants' dominant hemisphere. Furthermore, high correlations between the overall test results showed that participants with high scores on the verbal tests tended to achieve equally high scores on the non-verbal IQ test, suggesting that verbal and non-verbal abilities draw on similar mental resources. The analysis of game choice identified visuospatial skills as an important asset in language learning. Together with the evidence from previous studies in the field of SLA research, these findings indicate that cognitive mechanisms underpinning language in our brains can have a measurable influence on second language learning success. Yet, while these influences can be observed and measured, their nature and connections amongst themselves remain largely unexplained. Avenues for further research in these and related issues may include research on the relation between atypical speech representation and speech impairments, the search for a safe, non-invasive, and affordable measurement for cerebral dominance, and research on game preferences in relation to language aptitude.

References

Adams, E. A. (1952). *Analysis of raven's matrices scores: Preliminary report*. Surrey, UK: Surrey Educational Research Association.

Armour, J., Davison, A., & McManus, I. C. (2014). Genome-wide association study of handedness excludes simple genetic models. *Heredity (Edinb.), 112*(3), 221–225. https://doi.org/10.1038/hdy.2013.93

Beimer, N. J., Buchtel, H. A., & Glynn, S. M. (2015). One centre's experience with complications during the Wada test. *Epilepsia, 56*(8), 110–113. https://doi.org/10.1111/epi.13046

Bernat, E., & Lloy, R. (2007). Exploring the gender effect on EFL learners' beliefs about language learning. *Australian Journal of Educational & Developmental Psychology, 7*, 79–91. https://doi.org/10.21512/lc.v8i2.453

Biedrón, A., & Pawlak, M. (2016). New conceptualizations of linguistic giftedness. *Language Teaching, 49*(2), 151–185. https://doi.org/10.1017/S0261444815000439

Binder, J. R. (2011). Functional MRI is a valid non-invasive alternative to wada testing. *Epilepsy & Behaviour, 20*(2), 214–222. https://doi.org/10.1016/j.yebeh.2010.08.004

Boyle, J. P. (1987). Sex differences in listening vocabulary. *Language Learning, 37*(2), 273–284.

Briscoe, J., Chilvers, R., Baldeweg, T., & Skuse, D. (2012). A specific cognitive deficit within semantic cognition across a multi-generation family. *Proceedings of the Royal Society B, 279*(1742), 3652–3661. https://doi.org/10.1098/rspb.2012.0894

Carroll, J. (1981). Twenty-five years of research in foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.

Carroll, J. B., & Stanley, M. (1959). *Modern language aptitude test (MLAT)*. San Antonio, TX: Psychological Corporation.

Christiner, M., & Reiterer, S. M. (2015). A Mozart is not a Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience, 9*, 482. https://doi.org/10.3389/fnhum.2015.00482

Corballis, M. C. (2014). Left brain, right brain: Facts and fantasies. *PLoS Biology, 12*(1). https://doi.org/10.1371/journal.pbio.1001767

Corey, D., Hurley, M., & Foundas, A. (2001). Right and left handedness defined: A multivariate approach using hand preference and hand performance measures. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 14*(3), 144–152.

Dediu, D. (2008). The role of genetic biases in shaping the correlations between languages and genes. *Journal of Theoretical Biology, 254*(2). https://doi.org/10.1016/j.jtbi.2008.05.028

Dehaene-Lambertz, G., Dahaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science, 298*(5600), 2013–2015. https://doi.org/10.1126/science.1077066

DeKeyser, R. (2012). Interaction between individual differences, treatment, and structures in SLA. *A Journal of Research in Language Studies, 62*(2), 189–200. https://doi.org/10.1111/j.1467-9922.2012.00712.x

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell.

Dorsaint-Pierre, R., Penhune, V. B., Watkins, K. E., Neelin, P., Lercch, J., Bouffard, M., & Zatorre, R. J. (2006). Asymmetries of the planum temporale and Heschl's gyrus: Relationship to language lateralization. *Brain, 129*, 1164–1176. https://doi.org/10.1093/brain/awl055

Dronkers, N. F., Plaisant, O., Iba-Zizen, M. T., & Cabanis, E. A. (2007). Paul Broca's historic case: High resolution MR imaging of the brains of Leborgne and Lelong. *A Journal of Neurology, 130*, 1432–1441. https://doi.org/10.1093/brain/awm042

Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York: Random House.

Golestani, N. (2012). Brain structural correlates of individual differences at low to high-levels of the language processing hierarchy: A review of new approaches to imaging research. *International Journal of Bilingualism, 18*(1), 6–34. https://doi.org/10.1177/136700691245658

Golestani, N., Paus, T., & Zatorre, R. J. (2002). Anatomical correlates of learning novel speech sounds. *Neuron, 35*, 997–1010. https://doi.org/10.1016/S0896-6273(02)00862-0

Goodwin, R. S., & Michel, G. F. (1981). Head orientation position during birth and in infant neonatal period, and hand preference at nineteen weeks. *Child Development, 52*(3), 819–826.

Groen, M. A., Whitehouse, A. J., Badcock, N. A., & Bishop, D. V. (2013). Associations between handedness and cerebral lateralization for language: A comparison of three measures in children. *PLoS One, 8*(5), e64876. https://doi.org/10.1371/journal.pone.0064876

Hepper, P. G., Shahidullah, S., & White, R. (1991). Handedness in the human foetus. *Neuropsychologia, 29*, 1101–1111.

Johnston, D., Nicholls, M., Shah, M., & Shields, M. (2009). Nature's experiment? Handedness and early childhood development. *Demography, 46*(2), 281–301. https://doi.org/10.1353/dem.0.0053

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., & Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain: A Journal of Neurology, 123*(12), 2512–2518.

Kratzmeier, H., & Horn, R. (1980). *Manual: Raven-Matrices test. Advanced progressive matrices*. Weinheim, Germany: Beltz Test.

Language Learning and Testing Foundation. (2014). *Modern language aptitude test*. http://lltf.net/aptitude-tests/language -aptitude-tests/modern-language-aptitude-test-2/. Accessed 13 May 2016.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*, 251–285.

Meara, P. (2005). *LLAMA language aptitude tests the manual*. Swansea, UK: Lognostics.

Mercer, S. (2012). Dispelling the myth of the natural-born linguist. *ELT Journal, 66*(1), 22–29. https://doi.org/10.1093/elt/ccr022

Papadatou-Pastou, M. (2011). Handedness and language laterization: Why are we right-handed and left-brained? *Hellenic Journal of Psychology, 8*, 248–265.

Pelletier, I., Sauerwein, H. C., Lepore, F., Saint-Amour, D., & Lassonde, M. (2007). Non-invasive alternatives to the Wada test in the presurgical evaluation of language and memory functions in epilepsy patients. *Epileptic Disorders, 9*(2), 111–126. https://doi.org/10.1684/epd.2007.0109

Propper, R. E., O'Donnell, L. J., Whalen, S., Tie, Y., Norton, I. H., Suarez, R., & Golby, A. J. (2010). A combined fMRI and DTI examination of functional language lateralization and arcuate fasciculus structure: Effects of degree versus direction of hand preference. *Brain and Cognition, 73*(2), 85–92. https://doi.org/10.1016/j.bandc.2010.03.004

Pujol, J., Deus, J., Losilla, J. M., & Capdevilla, A. (1999). Cerebral lateralization of language in normal left-handed people studied by functional MRI. *Neurology, 52*(5), 1038–1043.

Raven, J. (2000). The Raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology, 41*(1), 1–48.

Raven, J. C. (1941). Standardisation of progressive matrices, 1938. *British Journal of Medical Psychology, 19*(1), 137–150.

Reiterer, S. M. (2009). Brain and language talent: A synopsis. In G. Dogil & S. M. Reiterer (Eds.), *Language talent and brain activity* (pp. 155–192). Berlin, Germany: Mouton de Gruyter.

Robinson, P. (2007). Aptitude, abilities, contexts, practices. In R. DeKyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 256–286). New York: Cambridge University Press.

Saville-Troike, M. (2006). *Introducing second language acquisition*. Cambridge, UK: Cambridge University Press.

Scerri, T. S., Brandler, W. M., Paracchini, S., Moris, A. P., & Ring, S. M. (2011). PCSK6 is associated with handedness in individuals with dyslexia. *Human Molecular Genetics, 20*(3), 608–614. https://doi.org/10.1093/hmg/ddq475

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.

Sparks, R., & Ganschow, L. (2001). Aptitude for learning a foreign language. *Annual Review of Applied Linguistics, 21*, 90–111. https://doi.org/10.1017/S026719050100006X

Sternberg, R. (2002). The theory of successful intelligence and its implications for language aptitude testing. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 13–44). Amsterdam: Benjamins.

Szaflarski, J., Holland, S. K., Schmithorst, V. J., & Byars, A. W. (2006). An fMRI study of language laterization in children and adults. *Human Brain Mapping, 27*(3), 202–212. https://doi.org/10.1002/hbm.20177

Whitehouse, A. J., & Bishop, D. (2009). Hemispheric division of function is the result of independent probabilistic biases. *Neuropsychologia, 47*(8–9), 1938–1943. https://doi.org/10.1016/j.neuropsychologia.2009.03.005

Willems, R. M., Toni, I., Hagoort, P., & Casasanto, D. (2009). Body-specific motor imagery of hand actions: Neural evidence from right- and left-handers. *Frontiers in Human Neuroscience, 3*(39). https://doi.org/10.3389/neuro.09.039.2009

Witkin, H. (1962). *Psychological differentiation*. New York: Wiley.

Xiang, H., Dediu, D., Roberts, L., Norris, D., & Hagoort, P. (2012). The structural connectivity underpinning language aptitude, working memory, and IQ in the perisylvian language network. *Language Learning, 62*(2), 110–130. https://doi.org/10.1111/j.1467-9922.2012.00708.x

The Impact of Speaking a Tone Language on Music Aptitude

Niloufar Saraei

Abstract A rapidly growing body of scientific research suggests that there is a close relationship between music and language in that training in one can influence the other. Recent studies show that the mechanisms involved in processing music and language can improve our understanding of each domain. Nevertheless, most of the studies have merely focused on the positive effects of music on language processing. Despite its significance, evidence for positive effects of language on music seems to be rare. The main purpose of this study is to investigate the correlation between speaking a tone language and musical aptitude. It is hypothesized that speaking a tone language has a positive effect on music perception and that musicians have a better capability in pitch discrimination in comparison to non-musicians. A sample of 40 participants was recruited and included four groups: tone-language speaking musicians/non-musicians, and native speakers of non-tonal languages with/without musical knowledge. Using Gordon's musical aptitude test, the positive effects of having a tone-language background, as well as having advanced musical knowledge on the perception of musical pitch, were investigated. The results showed that tone-language speaking non-musicians had a better ability in pitch discrimination than the non-musicians who were natives of non-tonal languages. Furthermore, the findings indicated that academically trained musicians outperformed non-tonal language speaking non-musicians in the presented music aptitude test. This leads us to the conclusion that having musical knowledge and a tone-language background can both have positive effects on music perception.

1 Introduction

A considerable number of studies confirm the strong correlation between music and language (e.g. Bidelman, Gandour, & Krishnan, 2011a; Koelsch et al., 2002; Patel, 2008; Slevc, Rosenberg, & Patel, 2009). In principle, language and music follow a

N. Saraei (✉)
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_10

similar brain structure. According to Besson and Schön (2001), both music and language are "rulebased systems composed of basic elements (phonemes, words, notes and chords) that are combined into higher-order structures (musical phrases and sentences, themes and topics) through the rules of harmony and syntax" (Besson & Schön, 2001, p. 235). Fonseca-Mora, Toscano-Fuentes, and Wermke (2011) suggest that language and music are two capacities with a mutual evolutionary history. Similarly, Brown (2000) states that:

> the musilanguage stage in evolution […] was neither linguistic nor musical but […] embodied the shared features of modern day music and language, so that evolutionary divergence led to the formation of two distinct and specialized functions with retention of the shared features conferred onto them by the joint precursor. (p. 277)

One of the most important parameters that is shared by music and language is 'pitch'. While the structure in music is based on a hierarchical arrangement of pitch, in language the structure relies on the hierarchical arrangement of morphemes, words, and phrases (Bidelman, Gandour, & Krishnan, 2011a; McDermott & Hauser, 2005). In comparison to musical tones, tone languages offer a great chance for studying the linguistic use of pitch, for in these languages different tones will change the lexical meaning of the words. An interesting example in Mandarin would be *mā, má, mǎ, mà*, (妈麻马骂) which means four completely different things – 'mother', 'hemp', 'horse' and 'scold'- when pronounced with different tones (Xu, 1997). Hence, the sentence *má mā mà mǎ* (麻妈骂马) could mean something like 'the hemp's mother scolds the horse', depending on its pronunciation. Phonologically, the four tones have the values of High-Level (as in mā), Mid-Rising (as in má), Falling-Rising (as in mǎ) and High-Falling (as in mà) (Xu, 1997). The difference in the meaning of these words is not related to the stress of the words but to their tone. In comparison to tone-languages, in a non-tonal language pronouncing a word with various pitch contours will not change the lexical meaning of that word (Duanmu, 2004). Considering the above example, it would be reasonable to assume that speakers of a tone language might have better perception in determining musical tones as well.

The present study attempts to investigate whether speaking a tone language can have a positive transfer effect on perceiving musical pitch. Gordon's music aptitude test was used to discover if tone language speakers obtain better results in this test relative to the speakers of non-tonal languages. Additionally, this study aims to investigate whether musically trained individuals have a greater capability to perceive musical pitches and rhythms compared to non-musically trained individuals without a tonal language background. This is because academically trained musicians are required to pass various ear-training courses, meaning that their ears are trained to recognize musical intervals and rhythms. In pursuance of these aims, this paper is concerned with two research questions:

1. Do Mandarin native speakers perform better at musicality tests than non-tonal language speakers?
2. Do trained musicians outperform non-tonal language speaking non-musicians because their ears are musically trained?

These questions will be addressed after the next section which will provide essential background information.

2 Tone Languages and Musical Ability

Only a limited number of studies have reported the positive effects of language on music. Most of the related research has focused on the relationship between tone languages and absolute pitch (e.g. Deutsch, Henthorn, Marvin, & Xu, 2006; Lee & Lee, 2010). However, a handful of studies have already investigated the correlation between speaking a tone language and musical perception, which will be discussed shortly.

In 2011, Bidelman, Gandour and Krishnan investigated the cross-domain transfer effects of pitch experience (music and language) on the auditory brainstem (Bidelman, Gandour, & Krishnan, 2011a). Three groups of participants, including native speakers of Mandarin Chinese, native speakers of English with advanced musical knowledge, and native speakers of English without musical training, were recruited, and their brainstem frequency-following response (FFR) was recorded. The choice of participants in this study was quite careful as all the participants were matched in age and education, were all right-handed, and showed normal hearing sensitivity. Using the autocorrelation algorithms, both the pitch tracking accuracy of the participants, as well as their pitch strength responses from the brainstem, were calculated. The results showed that "experience-dependent neural mechanisms for pitch representation at the brainstem level, as reflected in pitch-tracking accuracy and pitch strength, are more sensitive in Chinese and amateur musicians as compared to non-musicians across domain" (Bidelman et al., 2011a, pp. 429–430).

Shortly after conducting the aforementioned study, the authors carried out further research in order to determine whether the superiority of tone-language speakers in pitch encoding over musically untrained English speakers, as observed in the previous study, had any beneficial effect on the perception of musical pitch (Bidelman, Gandour, & Krishnan, 2011b). Hence, they investigated the brainstem representations and perceptual distinction of tuned and detuned musical notes in tone-language speaking non-musicians vs. English-speaking musicians/non-musicians. The results of the study demonstrated that musicians and tone-language speakers had stronger brainstem representation for musically relevant pitch compared to English-speaking non-musicians without pitch experience. However, the surprising part was that despite the higher brainstem encoding in pitch discrimination in tone language speakers, this neural improvement did not necessarily lead to perceptual benefits as it did in the group of musicians. In other words, although both musicians and tone-language speakers outperformed their English-speaking non-musician peers in neural representations for musical stimuli, only musicians could use this information efficiently at a perceptual level.

In a similar study conducted in 2013, Bidelman, Hutka and Moreno compared the performance of tone-language speakers to that of English speaking musicians/non-musicians regarding their basic auditory and music perception, as well as general cognitive abilities (Bidelman, Hutka, & Moreno, 2013). Fifty-four participants were recruited for the study including three groups of English-speaking musicians, English-speaking non-musicians, and Cantonese speakers with minimal or no musical knowledge. Much like the previous studies, none of the English-speaking participants had

had any exposure to a tone language so far. The participants were tested on basic auditory tasks as well as complex music perception. The findings indicated that both musicians and tone-language speakers were superior to non-musicians in pitch discrimination as well as in auditory perceptual performance (Bidelman et al., 2013). More specifically, the results demonstrated that, while musicians had enhanced performance in comparison to other groups, tone-language speakers were also superior to their English-speaking non-musician peers in all measures of pitch discrimination and music perception. Hence, the researchers concluded that both tone-language experience and musicianship are associated with superior music perceptual abilities. Besides, musicians and tone language speakers outperformed non-musicians in working memory capacity, suggesting that having advanced musical knowledge, as well as a tone-language background, are both linked with enhanced general cognitive abilities in addition to basic perceptual enhancements.

As can be seen, the existing studies on the relationship between speaking a tone language and music aptitude are rare and contradictory. Since it seems that there is not yet enough evidence to prove the possible transfer effects of language on music, conducting further research in this area seems to be necessary.

Before examining the current study, the historical background of Gordon's music aptitude test will be discussed. As will be seen in the next subsection, some of Gordon's statements about an individual's music aptitude might be in conflict with the findings of the previously mentioned studies.

2.1 The Historical Background of Gordon's MAP

In 1919, Carl Seashore published 'Seashore Measures of Musical Talents' in which he discussed measuring musical abilities. It was the first time that an author provided a tool for measuring musical talents (Walters, 1991). After Seashores' work, numerous books and articles, as well as new music aptitude tests, were published in order to investigate "whether music aptitude was a product of nature or a product of nurture" (Walters, 1991, p. 66).

In 1965, Edwin Gordon published the Musical Aptitude Profile (MAP) as a result of 8 years of research, which is a musical aptitude test consisting of seven subtests (Gordon, 1989). Gordon's music aptitude test is believed to be the world standard in music aptitude testing. As Walters quotes, Gordon's MAP "possesses the highest standards for reliability and validity ever obtained by the author of a music aptitude test" (Walters, 1991, p. 66).

Gordon himself states:

> Over the past 25 years, more than 100 studies have been undertaken to directly investigate the reliability and validity of the Musical Aptitude Profile. […] As a result of all of those studies, the reliability and validity of the Musical Aptitude Profile has become well-established throughout the world. (1989, p. 4)

Gordon (1984) believes that no one is born without music talent. Every person is skilled in music to some extent. He claims that music talent is a product of innate potential as well as environmental impacts, and that the level of music with which a person is born can be nurtured through good conditions of environment until age 9. Gordon adds:

> […] unless innate potential is fostered before age nine (the sooner the better), environmental influences will no longer have an appreciable, if any, effect on innate potential. After approximately age nine, music aptitude is no longer developmental. It becomes stabilized. The impact of appropriate informal guidance in music at the earliest possible age is enormous. (Gordon, 1984, p. xiv)

Gordon also believes that the degree of musical aptitude between individuals can differ significantly. According to Gordon (1984), more than two-thirds of people in the world have an average music aptitude. The rest are either naturally very talented musicians or innately very untalented in music. Gordon adds, only one in one thousand people have an extremely high music aptitude with the "potential to achieve as a genius" (1984, p. 46).

Gordon (1984) continues that the important distinction between 'music achievement' and 'music aptitude' should be taken into consideration. Students with great achievements in music must necessarily have a high aptitude in music; however, students with low musical achievement may still have a high level of music aptitude. It means that the students with high musical aptitude but low achievements may not have been able to fulfill their potential due to different issues such as lack of instruction or motivation.

According to Gordon, music aptitude is an innate capacity which can only be nurtured until age 9; after this age, environmental influences (e.g. musical training) cannot affect it anymore. Nevertheless, the literature showed that musically trained individuals showed obvious superiority to non-musicians in the music aptitude test, though not all of them had begun their musical training prior to age 9.

2.2 Hypotheses

In the present study, using Gordon's Music Aptitude Profile, the music aptitude of four groups of participants (tone-language speaking musicians/non-musicians, and non-tonal language speaking musicians/non-musicians) is assessed to determine if tone-language speakers and trained musicians display an enhanced performance level. It is expected that speaking a tone-language results in an enhanced perception of musical pitch. Moreover, as most of the musicians in our study have announced their age of onset for musical training to be after age 9, it is expected that advanced musicians, regardless of the onset age of their musical training, have a better understanding of musical tones. Hence, we assess the music aptitude of both musicians and tone-language speakers and compare them to non-musicians to see if native Mandarin speakers obtain similar results to musicians in the music aptitude test. The following two hypotheses can therefore be formulated: H1 assumes that Mandarin native speakers perform better at Gordon's MAP than non-tonal language speakers. H2 expects that trained musicians outperform non-tonal language speaking non-musicians at Gordon's MAP.

3 Methods

3.1 Participants

The following four groups of participants were selected for the present study:

1. Tone-language speaking musicians, i.e., people whose mother tongue was Mandarin Chinese and who were professional musicians or music students, hereafter referred to as Chinese musicians (CM).
2. Tone-language speaking non-musicians, i.e., native speakers of Chinese without musical knowledge, hereafter referred to as Chinese non-musicians (CNM).
3. Non-tonal language speaking musicians, meaning professional musicians with various nationalities whose native language was not a tone-language, hereafter referred to as international musicians (IM).
4. Non-tonal language speaking non-musicians, i.e., people with various nationalities who had little or no musical knowledge and were native speakers of non-tonal languages, hereafter referred to as international non-musicians (INM).

The total number of the participants was 40, with 10 participants in each group. The mean age of the participants was 24.40, with the youngest being 19 and the oldest being 39. All the Chinese participants were born and raised in China and were speakers of Mandarin Chinese. The international participants in both groups of musicians and non-musicians were selected from various countries including Austria, Hungary, Slovakia, Poland, Turkey, Iran and Egypt. The majority of the participants had not finished a university degree, which was to be expected due to their young age. The rest had Bachelor's or Master's degrees. The number of males and females in the sample was almost equal (females = 21, males = 19). Most of the participants spoke three languages including English, German (as most of them were residents of Austria) as well as their mother tongue, which was Mandarin Chinese for half of the participants. The rest of the participants had certain degrees of proficiency in two, four, or five languages ($M = 3.2$). However, it should be noted that there were diverse levels of proficiency for the languages spoken by different individuals, with most having an average knowledge in foreign languages. None of the international participants had any exposure to a tone-language and there was no bilingually raised participant in the survey.

A significant issue that needed to be taken into account was the definition of 'musicians'. In this study, musicians were defined as professional performers or singers with at least 7 years of continuous training in music on their particular field. The age of onset of musical training varied among the musicians in the range of 5–18. Many of the musically trained participants, especially in the group of international musicians, had started their music training after age 9. Furthermore, it might be beneficial to distinguish between musicians who are still active and those whose musical skills might be somewhat rusty, since many abilities are often lost if they are not harnessed. In the current study, all the musicians described themselves as quite active in music-making within the past 5 years. Such a definition of musicians is roughly similar to the definitions used in many previous studies in which the neurological effects of musical training have been investigated (e.g. Bidelman & Krishnan, 2010; Chandrasekaran, Krishnan, & Gandour, 2009; Cooper & Wang, 2012; Parbery-Clark, Skoe, & Kraus, 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007).

It should be noted that no distinction was made between classical, pop, rock, and other genres; however, most of the musicians were trained in classical western music. There are two reasons for this: On the one hand, the samples in this study were mostly music students in Vienna who are trained to become a part of professional orchestras and classical concerts. On the other hand, most Chinese musicians in Austria are attracted by the international reputation that Austria has in the world of classical music.

Those participants who were categorized as non-musicians had no more than 3 years of musical training throughout their lifetime and in most cases, they had not received music lessons within the past 3 years.

3.2 Instruments

Three categories comprise the results of the MAP. While two categories test tonal and rhythmical skills, one is composed of the total score. The participants were asked to listen to musical statements followed by a musical answer. They had to recognize if the musical answer was different from the musical statement and if this difference was a tonal or a rhythmical one. In fact, they had to choose one of the three options available in the answer sheet: Rhythmical change, tonal change, and no change. The test consisted of 30 tracks in total and the participants had to listen to each track only once. There was no case with a simultaneous change of both rhythm and tone.

Consequently, the resulting categories were 'tone', 'rhythm', and 'total'. The highest possible score in the Gordon test is 100; however, raw scores differ from the international scale with a maximum of 80. While general scores were useful for seeing differences and classifying those into categories such as 'high musical talent' or 'average musical aptitude', calculating and testing with raw scores was favored.

3.3 Procedures

Gordon's music aptitude test is an online test, however, participants were tested face-to-face in order to guarantee that the test procedures were followed properly, i.e., that tracks were only listened to a single time. To facilitate this method, questionnaires were printed out and distributed among the participants. They were then asked to listen to music tracks playing from a laptop or music player and to write their answers on the answer sheet. Using this technique, two or more participants could be tested at the same time. It took the participants approximately 30–35 min to complete the whole procedure (Gordon's MAP took 20 min, the explanations and basic questionnaire 10–15 min).

4 Results

The first step in analyzing the data was to check for normality of distribution. In order to find out if the obtained data is normally distributed, three tests were used. Firstly, normality of distribution was tested by using the Kolmogorov-Smirnov test. Examining the scores reached in the categories of tone, rhythm, and total scores, the Kolmogorov-Smirnov test reported an asymptotic significance of $p = 0.2$ for each of the categories. By creating histograms and Q-Q plots, the normality of distribution was further checked. The histograms of the tonal, rhythmical and total scores were very similar since the rhythmical and tonal aspects of musical talent are highly connected. Musical training involves both aspects and tone and rhythm always go hand in hand in music. This is why it also does not come as a surprise that the Q-Q plots of the tonal, rhythmical and total scores were highly similar.

Since all the three tests proved the normality of distribution, the data could be analyzed by parametric statistical tests. In the next step, an ANOVA (Analysis of variance) was used in order to check the mean scores of each group as well as the standard deviations. Looking at the means of the four groups, it can already be observed that there are definite differences in the average scores. With the mean of 67.60 points ($SD = 6.04$), the group of Chinese musicians (CM) achieved the highest scores in comparison to others.

With a slight difference, international musicians (IM) achieved the second highest mean of 67.40 points ($SD = 7.86$). Included in this group, one participant reached the highest possible score with a raw score of 80 points.

Not much different from the average score of the two aforementioned groups, the group of Chinese non-musicians (CNM) achieved the mean of 62 points ($SD = 6.37$).

It is obvious that there is not an immense difference between the results of the three aforementioned groups (CM, CNM and IM). However, the participants of INM group, who were neither tone-language speakers nor had musical knowledge, have the mean score of 47.20 points ($SD = 7.04$). This is much lower than the other three groups, and among them the lowest score was reached with 36 points.

Next, the Levene's test was used to assess the homogeneity of variances. This test provides an F statistic and a significance value. If the significance level is greater than 0.05, the group variances can be treated as equal. The p-value in our case is 0.800 which is much greater than 0.05; this confirms the homogeneity of variances.

The ANOVA shows a significant value of $p < 0.01$, which confirms that there is a significant difference between our groups somewhere. However, it does not make it clear which groups are significantly different from one another. Hence, the post-hoc test was performed in order to clarify the significant differences between specific groups. The outcome of the post-hoc test with multiple comparisons can be seen in Table 1 below.

In this case, the Tukey test was chosen to reveal the significant relationships between groups. The asterisks in the above chart indicate the significantly different groups. As can be seen, the INM group has a significant difference to all other groups. The existence of such a difference confirms the hypothesis that

Table 1 Post hoc test

Dependent variable: Gor_Total_raw		
Tukey HSD		
(I) groups	(J) groups	Mean difference (I–J)
CM	CNM	5.60000
	IM	.20000
	INM	20.40000*
CNM	CM	−5.60000
	IM	−5.40000
	INM	14.80000*
IM	CM	−.20000
	CNM	5.40000
	INM	20.20000*
INM	CM	−20.40000*
	CNM	−14.80000*
	IM	−20.20000*

*The mean difference is significant at the 0.05 level

tone-language speakers (who are not musically trained) have greater capabilities in pitch discrimination in comparison to the speakers of non-tonal languages (who also do not have musical knowledge).

Nevertheless, while musicians (in both groups of Chinese and international) resulted in a higher mean relative to non-musicians, the results did not show any significant difference between groups CNM and CM as well as groups CNM and IM.

The difference between the mean scores of the four groups is presented in Fig. 1 below.

5 Discussion

In an effort to determine the possible transfer effects of having a native tone-language background, as well as having advanced musical knowledge on music perception skills, a cross-sectional study of 40 adults, including four groups of Mandarin-speaking musicians/non-musicians, and non-tonal language speaking musicians/non-musicians was performed. Using Gordon's Music Aptitude Profile (MAP), music perception abilities of the participants were tested in order to investigate if there is any advantage in perceiving musical tones in tone-language speakers and professional musicians. The results showed that both trained musicians and native Mandarin speakers have superior performance in pitch discrimination and music perception, as compared to those participants without musical and tone language experience (international non-musicians). This result, in fact, corroborates the existence of a positive relationship between language and music domains and therefore concurs with other studies that argue for the neuronal association between

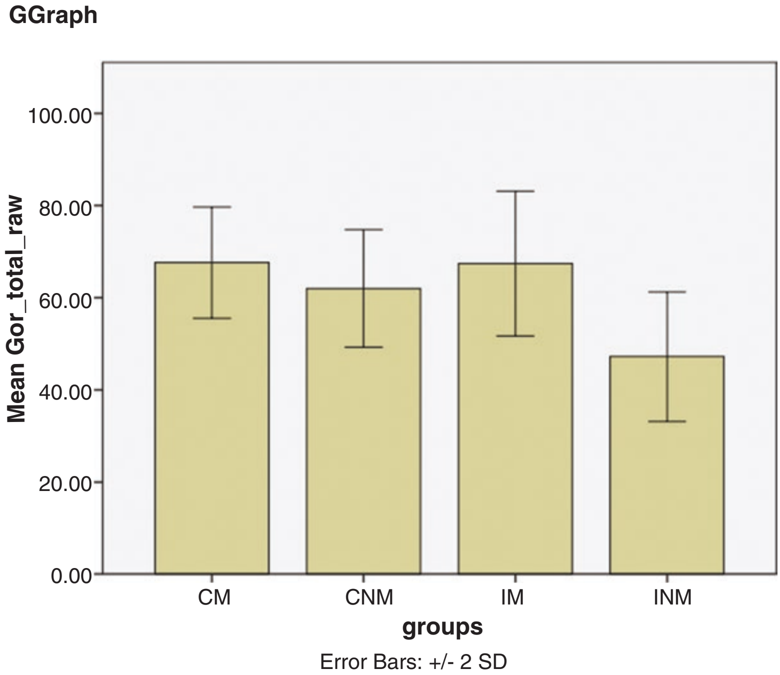


Fig. 1 Bar chart of the means

language and music (see Brown, 2000; Fonseca-Mora, Toscano-Fuentes, & Wermke, 2011; Besson & Schön, 2001; Patel, 2008).

The first hypothesis of this study, which assumed that Mandarin native speakers perform better at Gordon's MAP than non-tonal language speakers, was significantly supported. The results showed that tone-language speakers who were musically untrained obtained significantly better results in the musicality aptitude test relative to non-tonal language non-musicians. This result is broadly consistent with the findings of Bidelman, Hutka, and Moreno (2013) who found that tone-language expertise goes together with superior music perceptual abilities. As already stated in the literature review, the study compared the pitch discrimination and music perception abilities of three groups of English speaking musicians/non-musicians and tone language (Cantonese) speakers without musical training. The results indicated that trained musicians and tone-language speakers have superior auditory acuity and music perceptual abilities in comparison to English speaking non-musicians, suggesting that both tone language and musical expertise are associated with enhanced perceptual abilities for musical pitch.

In contrast, the results of our experiment seem to be inconsistent with the findings of Bidelman et al. (2011b), as their findings did not support superior auditory and music perceptual abilities for tone-language speakers. As mentioned before, they conducted a survey in 2011 (Bidelman et al., 2011a) in which they found out that both musicianship and tonal language experience might be associated with enhanced brainstem encoding of pitch-relevant information. By conducting further investigation, they tried to determine if this superiority in pitch encoding could have positive effects on perception of music (Bidelman et al., 2011b). Their findings, unlike ours, demonstrated superior perceptual advantages for musicians only. No perceptual benefits for music perception were observed in tone language speakers.

Bidelman et al. (2013) suggest that the 2011 study (Bidelman et al., 2011b) probably failed to prove the relationship between tone language experience and enhanced music perception as the Mandarin linguistic system contains exclusively curvilinear lexical tones, which are simply incompatible with musical pitch patterns, i.e., the auditory patterns of music domains may fall outside the scope of tone language domains. In other words, perception of complex musical patterns might require extra skills which are not possessed by tone-language speakers. This argument could also explain why musicians have the highest results in nearly all measures of pitch discrimination.

The second hypothesis of this study, which expected that trained musicians outperform non-tonal language speaking non-musicians at Gordon's MAP, was also supported. Indeed, the results of Gordon's music aptitude test obtained by musicians in both the Chinese and international groups were significantly higher than the results of international non-musicians. Moreover, professional musicians outperformed tone- language speaking non-musicians in the MAP test; however, the difference between their scores did not reach statistical significance.

The present data are therefore consistent with the data reported in the literature. As already stated, Bidelman et al. (2011a) reported better results for musicians and Mandarin speaking individuals in pitch-tracking accuracy and pitch strength, as compared to English speaking non-musicians. Yet, their result showed that the pitch strength in musicians was even greater than in Chinese listeners.

The findings of their next survey (Bidelman et al., 2011b) also revealed the superiority of musicians to Chinese and non-musicians in perception of musical pitch. Similar to our findings, Bidelman et al. (2013) also reported the enhanced performance of musicians and tone language speakers in auditory and pitch discrimination tasks, albeit with musicians outperforming the Chinese participants in most of the tasks (though the difference often failed to reach significance).

Considering the superiority of musicians to non-musicians in our study as well as in the literature, we could challenge Edwin Gordon's theory which claims that it is not possible to nurture musical talent from the age of nine onwards, as many of the musicians in these studies began their musical training after age 9. Yet, they significantly outperformed their non-musician counterparts in Gordon's music aptitude test.

When comparing our study to the existing literature, we must take into consideration one important issue; namely, the inclusion of the group of tone-language speaking musicians in this research, which is missing in the previous similar studies (Bidelman et al. 2011a, 2011b, 2013). Indeed, we included the group of Chinese musicians to assess if there is any additional perceptual advantage for this group. In other words, we wanted to examine the combined effects of tone language and musical experience. Our results showed that the scores obtained by the group of

Chinese musicians are quite similar to the scores of international musicians, i.e., the combination of the two experiences did not cause extra perceptual advantage for tone-language speaking musicians.

6 Conclusion

Our study was successful in that we could support both of our hypotheses which assumed that speaking a tone-language and being a musician increase musical ability as measured by Gordon's MAP test. However, as with every research, the current study has some limitations which should be taken into account when interpreting the results. First, Gordon's music aptitude test was an online test in which the participants had to register for the test via the invitation link. Since the participants were not observed while doing the test, one could not guarantee that they listened to each track only once. Evidently, listening more than once could have changed the results significantly.

To solve this problem, some of the students were tested in a different way. Questionnaires were printed out and the participants were asked to listen to music tracks playing from a laptop or music player and write their answers on the answer sheet. Using this technique, two or more participants could be tested at the same time. However, the disadvantage of this technique was that sometimes the environment was too noisy for the participants to concentrate well, and the music was not loud enough for the participants to hear.

Second, the music aptitude test was quite time-consuming. The participants had to listen to 30 music tracks as well as explanations and sample tracks. The explanation itself sometimes took a lot of time since some of the non-musician participants had absolutely no understanding of musical tone and rhythm. Hence, after listening to ten, or at maximum twenty tracks, most of the participants announced that they could not concentrate anymore and wanted to stop doing the test.

Additionally, one cannot guarantee that the superiority of Chinese non-musicians to their musician peers was totally due to their ability to speak a tone language. Many other factors could have been involved. For example, it could be possible that the superiority of musicians and tone-language speakers was simply due to their personal intelligence, biological endowment, or even genetics. Yet there are other factors, such as cultural influences, that cannot be measured easily. An example of such cultural factors would be the differences between the education systems in different countries. For instance, it could be that early music training in China's schools provides Chinese students with a better ability to perceive musical pitch. Further research is needed before definitive conclusions can be drawn regarding the relationship between tone-language background and music aptitude. What is more, increasing the number of participants will provide more reliable results, as a higher number of participants would clearly be a better representation of the population.

Another suggestion for further research would be to test the tone-language speaking participants and musicians on their language aptitude and to investigate whether there is any relation between their language aptitude and music perception. In other words, assessing the language aptitude of the participants in various areas such as phonetic memory and unintelligible speech imitation would help us determine if their musicality has further positive transfer effects on their language learning domain.

Finally, this topic can be extended immensely since there are many details to be considered in future studies, which means that this study created a valuable basis for further research regarding the positive transfer effects of language experience on musical abilities.

References

Besson, M., & Schön, D. (2001). Comparison between language and music. *Annals of the New York Academy of Sciences, 930*(1), 232–258. https://doi.org/10.1111/j.1749-6632.2001.tb05736.x

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011a). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*, 425–434. https://doi.org/10.1162/jocn.2009.21362

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition, 77*(1), 1–10. https://doi.org/10.1016/j.bandc.2011.07.006

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: Evidence for bidirectionality between the domains of language and music. *PLoS ONE, 8*(4), 1–11. https://doi.org/10.1371/journal.pone.0060676

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research, 1355*, 112–125. https://doi.org/10.1016/j.brainres.2010.07.100

Brown, S. (2000). The “musilanguage” model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271–300). Cambridge, MA: MIT Press.

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*(1), 1–9. https://doi.org/10.1016/j.bandl.2008.02.001

Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America, 131*(6), 4756–4769. https://doi.org/10.1121/1.4714355

Deutsch, D., Henthorn, T., Marvin, E., & Xu, H. (2006). Absolute pitch among American and Chinese conservatory students: Prevalence differences, and evidence for a speech-related critical period. *The Journal of the Acoustical Society of America, 19*(2), 719–722. https://doi.org/10.1121/1.2151799

Duanmu, S. (2004). Tone and non-tone languages: An alternative to language typology and parameters. *Language and Linguistics, 5*(4), 891–924.

Fonseca-Mora, C., Toscano-Fuentes, C., & Wermke, K. (2011). Melodies that help: The relation between language aptitude and musical intelligence. *Anglistik International Journal of English Studies, 22*(1), 101–118.

Gordon, E. E. (1984). *Learning sequences in music: A contemporary music learning theory*. Chicago: GIA Publications.

Gordon, E. E. (1989). Testing music aptitudes from preschool through college. *Gordon Institute for Music Learning Newsletter, 2*, 4–6.

Koelsch, S., Gunter, T. C., v Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical “language-network” serves the processing of music. *NeuroImage, 17*(2), 956–966. https://doi.org/10.1006/nimg.2002.115

Lee, C., & Lee, Y. (2010). Perception of musical pitch and lexical tones by Mandarin-speaking musicians. *The Journal of the Acoustical Society of America, 127*(1), 481–490. https://doi.org/10.1121/1.3266683

McDermott, J., & Hauser, M. (2005). The origins of music: Innateness, uniqueness, and evolution. *Music Perception: An Interdisciplinary Journal, 23*(1), 29–59. https://doi.org/10.1525/mp.2005.23.1.29

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 29*(45), 14100–14107. https://doi.org/10.1523/JNEUROSCI.3256-09.2009

Patel, A. D. (2008). *Music, language, and the brain*. New York: Oxford University Press.

Slevc, L., Rosenberg, J., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review, 16*(2), 374–381. https://doi.org/10.3758/16.2.374

Walters, D. L. (1991). Edwin Gordon’s music aptitude work. *The Quarterly Journal of Music Teaching and Learning, 2*(1–2), 65–72.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*(4), 420–422. https://doi.org/10.1038/nn1872

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics, 25*(1), 61–83.

Making Music and Learning Languages – Musicality and Grammar Aptitude

Daniel Malzer

Abstract Research on correlations between musicality and language aptitude have been predominantly investigating the phonetic aspect of language processing. The current state of research suggests a strong and stable link between musicality and receptive language abilities, such as recognition of sounds, intonation and stress patterns, as well as productive skills. Relatively fewer studies have explored relations of musicality and grammar aptitude, despite neurological studies highlighting similar brain regions involved in the processing of musical, especially rhythmic, as well as grammatical patterns. This paper thus aims to investigate if musical training and musicality does indeed relate to grammatical skills. It is hypothesised that extensive musical training does not only impact the musical ear but also the ability to de- and encode structures, as well as the capacity to recognise and retain complex sequences. These specific skills are widely recognised to be involved in the acquisition of novel grammar. Research was conducted by testing a sample of 25 participants, which was split into two groups, musicians and non-musicians. Musicality of all participants was assessed and a grammar achievement test was issued. The results suggest a strong correlation between musical training, musicality and grammatical aptitude.

1 Introduction

> What Plato says about all the fine arts as fostering learning can be applied to music in particular. He speaks of these arts as preparing the mind for understanding by providing a cultural formation. […] For Plato music directly touches the emotions and remotely prepares the intellect for learning, so that this end which refers to the intellectual life is consequent upon its effect in the moral order. (Schoen-Nazarro, 1978, p. 265)

Musical abilities, a good ear, and participating in social exchanges through music have been related to general intelligence and to superior education since the Greek

D. Malzer (✉)
Department of English and American Studies, University of Vienna, Vienna, Austria

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_11

and Roman empires. Scholars were not only trained in disciplines such as grammar, rhetoric or mathematics but also, within the framework of the *artes liberales*, music.

In recent years, research has extensively explored connections of intelligence, musical aptitude and the talent to acquire new languages. These studies all share a common goal: to determine if, and to what degree, musicality is affecting, causing, or correlating with personal traits, such as the aforementioned talent to acquire new languages. The vast majority of studies investigating the connection of language aptitude and musicality have dealt almost exclusively with the auditory segment of language processing: pronunciation, differentiation of stress, sounds or intonation patterns. Generally, it can be concluded that musicality does indeed correlate strongly with phonetic processing of languages.

However, research into possible correlations between musicality and different fields of language acquisition has been relatively limited thus far. In a recent conference paper, Kalcheva and Fonseca-Mora underlined this apparent gap, pointing out that only a "few studies [have been] contributing to the relationship and influence of music (…) on grammar achievement" (2017, p. 391). The paper at hand attempts to add to this under-represented question: does musical aptitude correlate with grammatical pattern recognition and reproduction? Or, put in other words, do the observed correlations of language aptitude and musicality also apply to the processing of written words and sentences? Numerous studies suggest that similar brain regions are involved in the processing of musical as well as language syntax. Additionally, the long standing, but highly disputed, bootstrapping theory also suggests a strong link between the recognition of acoustic features and the development of syntactical awareness in first language acquisition of infants (see for example the critical analysis of Fernald & Mcroberts, 1996). Clearly, having received musical training does foster and develop a wide array of skills and abilities, especially in an interactional setting with other musicians. Adaptability, recognition and reproduction of rhythmic and melodic patterns and retention of complex sequences may all have an impact on the processing of grammatical structures of a language.

The conducted small scale study does indeed suggest that musicians with musical training and active participation in a musical setting do display superior grammatical aptitude. The study also indicates that sub-skills of musical aptitude including the processing of patterns such as melodic and rhythmic progressions, correlate more significantly with grammatical de- and encoding as opposed to phonetic processing of pitch and tempo.

The first section of the paper at hand aims to establish a common ground on the definitions of language aptitude, musicality and grammar aptitude, including a conclusive overview of the current state of research. Subsequently, the methodology and the results of the study will be presented, followed by the discussion section attempting to relate the results to the state of research.

1.1 Language Aptitude

Human beings differ greatly in the effort required to learn a second language. Dörnyei emphasises that these individual differences “refer to dimensions of enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree “(2005, p. 4). These deviations from the average are regularly referred to as talent, trait, innate abilities or qualities.

The explanation of why individuals vary can be divided into internal and external factors (Jilka, 2009, p. 1). On the one hand, internal factors encompass biological and genetic factors such as intelligence, innate personality traits relating to motivation, or empathy and aptitude (Jilka, 2009, p. 1). External factors range from socio-economic circumstances and culture-specific environments to teaching and learning strategies and methods. Methodology and strategies for teaching and learning constitute the major focus of the last decades of research into second language acquisition (Dörnyei & Skehan, 2003, 593).

The main questions leading the field of research are if such an intrinsic talent for language aptitude can be measured, if it can predict learning success effectively, and how it relates to external factors such as context, methodology and sociological backgrounds (see Dörnyei & Skehan, 2003, p. 591). According to Carroll, language aptitude encompasses four components: the ability decode and encode unfamiliar sounds, the ability to identify grammatical functions in larger segments of language, the extraction of syntactic and morphological patterns and the application in new chunks of language and, finally, the generation of an associative memory, linking vocabulary between L1 and L2 (Carroll 1962, qtd. in Dörnyei & Skehan, 2003, p. 592). The Modern Language Aptitude Test, devised by Carroll consequently defines language aptitude through four complementary abilities: “phonetic coding ability, grammatical sensitivity, rote learning ability, inductive language learning ability” (Dörnyei, 2005, p. 39–40).

Apart from the Modern Language Aptitude Test, few other aptitude tests have gained as much influence on general research into aptitude. Noteworthy are the Pimsleur Language Aptitude Battery (Pimsleur, 1966) and, more recently, the CANAL-F (Cognitive Ability for Novelty in Acquisition of Language – foreign) battery devised by Grigorenko et al. (2000). While Pimsleur‘s test battery is “quite similar to Carroll‘s MLAT “(Dörnyei & Skehan, 2003, p. 594), the Canal-F test is based on the theory that the acquisition of a language is related to general knowledge acquisition (Grigorenko et al., 2000, p. 392). This approach emphasises the central ability to cope with novelty and ambiguity in the processing of new information of an unknown language (Grigorenko et al., 2000, p. 392). Finally, the LLAMA test battery (Meara, 2005) has seen considerable use in recent years. This aptitude test battery is “loosely based” (Meara, 2005, p. 2) on the MLAT by Carroll and Sapron, using mostly picture stimuli to negate the influencing factor of differing L1 and L2 backgrounds.

Despite the widespread use of aptitude tests, achieving reliable results is rather difficult. As Jilka points out, the difficulty lies within the differentiation between talent and external factors: “accordingly, individual test tasks should be defined and constructed in such a way that the targeted abilities are indeed investigated” (Jilka, 2009, p. 8). Jilka refers to the general concept of construct validity and reliability, a highly debated topic in teaching methodology in regard of testing and assessment (see for example Brown and Abeywickrama 2010, p. 30, or Hughes, 2003, p. 26). Jilka, furthermore, advocates for the control of as many of external factors as possible, to “get at the core of ‘talent’” (Jilka, 2009, p. 9). Hence, in order to exclude experience, practice, and L1 proficiency (that can obviously vary greatly despite it being the native language) Jilka proposes the use of artificial or unknown languages and “a large homogeneous group of the same age and ‘learning career’” (Jilka, 2009, p. 9).

Despite these inherent limitations and difficulties of testing aptitude, Dörnyei and Skehan highlight the importance of research, as, aside from age of onset, “language aptitude and motivation have generated the most consistent predictors of second language learning success”(2003, p. 589).

1.2 Musicality

Honing et al. define musicality as “as a natural, spontaneously developing trait based on and constrained by biology and cognition” (2015, p. 1). Thus, while it is unclear to which extent musicality is an innate talent or an acquired and trained skill there is no doubt that musicality encompasses “many different components, ranging from perceptual capacities for detecting pitch and rhythm, as well as motor capacities, to emotional/theory of mind capacities for anticipating an audience‘s reaction” (Marcus, 2012, p. 501). Moreover, despite possible biological predispositions, attaining musical proficiency is “significantly correlated with amount of practice” (Marcus, 2012, p. 503), or as Bermudez et al. emphasise: the “intensive training and practice involved in achieving high levels of musicianship place extraordinary demands on many of the mind‘s most critical faculties”(2009, p. 1583).

It is not only the de- and encoding of musical information in the reception and production that requires training and practice. Especially the interaction with other musicians involves pattern recognition and retention as well as improvisation as indicated by Koelsch (2005, p. 207). Volz emphasises the complex factors that intertwine when improvising and writing music (2005, p. 50). Kraus and Chandrasekaran support this notion in their experiments, ascertaining that

> [a]ctive engagement with music improves the ability to rapidly detect, sequence and encode sound patterns. Improved pattern detection enables the cortex to selectively enhance predictable features of the auditory signal at the level of the auditory brainstem (2010, p. 600).

In respect of anatomical characteristics, a multitude of studies observed structural differences within the brain of musicians in comparison to non-musicians.

Kraus and Chandrasekaran observed "increased neural activity (…) in the auditory cortex" of pianists while hearing piano music (2010, p. 599), while various voxel-based morphometries show increased grey matter density in the Broca's area of musicians (Sluming et al., 2002) as well as in the Herschl's gyrus and left inferior frontal gyrus (Gaser & Schlaug, 2003, James et al., 2014). Moreover, Maess et al. conducted a magnetoencephlaography, showing that "harmonically inappropriate chords activated Broca's area and its right-hemisphere homologue" (2001, p. 543, similarly Marques et al., 2007). Generally, it is agreed upon that musical training and expertise can be traced through differing brain structures. Strait and Kraus presume that these changes are caused by the extraordinary demands of processing music:

> [d]ue to its multisensory nature, attentional demands, complex sound structure, rhythmic organization and reliance on rapid audio-motor feedback, music is a powerful tool for shaping neuronal structure and function (Strait & Kraus, 2011, p. 141).

1.3 Studies on Language Aptitude and Musicality

The main line of argument accompanying studies on language aptitude and musicality is the concept that musical practice trains the brain to be more perceptive. Hence, this increased auditory fitness is likely to affect not only the perception of sounds but also the production (see Kraus & Chandrasekaran, 2010, p. 599). Various studies have highlighted increased language aptitude of musicians compared to non-musicians in relation to receptive phonetic skills: increased pitch processing (Marques et al., 2007); better discrimination of tonal and segmental variations (Slevc & Miyake, 2006, p. 679; Marie et al., 2011); and increased phonetic awareness in distinguishing between phonemes and intonation (Fonseca-Mora, Toscano-Fuentes & Wermke, 2011, p. 105; Pastuszek-Lipinska, 2004, p. 68). Comparatively fewer studies also tested and observed increased productive abilities of musicians: improved pronunciation (Milovanova et al., 2008), or better performance on language imitation tasks (Christiner & Reiterer, 2015). Additionally, Christiner and Reiterer observed better results of vocalists compared to instrumentalists on a language imitation task of an unknown language (2015). These results are supported by a rather comprehensive study among 128 Chinese college students, again showing strong correlations between musical aptitude and suprasegmental production in a foreign language (Pei et al., 2016, p.19). In their extensive literature review, Chobert and Besson conclude:

> Taken together, these results show that musicianship facilitates the learning of non-native supra-segmental and segmental contrasts defined by acoustical features (e.g., pitch and duration) and improves categorical perception. It may be that musical expertise refines the auditory perceptive system (bottom-up facilitation), but it may also be that years of intensive musical practice exert top-down facilitatory influences on auditory processing. (2013, p. 928)

1.4 Grammar and Syntax Processing

Brown and Abeywickrama (2010, p. 294) define grammatical competence as knowledge of grammatical forms as structure of the language, as the meaning of these forms and, finally, as the pragmatic meaning in its corresponding context. Concerning the form, the authors emphasise that "form is both morphology, or how words are formed, and syntax, how words are strung together" (Brown & Abeywickrama, 2010, p. 294). These conceptual categories are based on the works of James Purpura who, in great detail, elaborated on the specific sub-categories of each area (Purpura, 2004, p. 91). Specifically for the sentential level, Purpura refers to the segmental and lexical forms, orthographic, syntactic and morphological features and irregularities, as well as word formation and morphosyntactic forms and affixes (Purpura, 2004, p. 91). However, Purpura additionally considers prosody and correspondence of sound and spelling as well as phonetic features as part of grammatical processing (2004, p. 91). Thus, following this paradigm, the process of de- and encoding of sounds and the transfer to the written word is all entrenched in the wider area of grammatical knowledge. Flöel et al. on the other hand, regard the extraction of rule-based information as the core and intrinsic requirement for the acquisition of grammar (2009, p. 1974). Hence, grammatical learning involves predominantly rule extraction to create and assess knowledge (Flöel et al., 2009, p. 1975). In a similar vein, Kepinska et al. emphasise the analytical ability as the dominant component while acquiring novel grammar (2016, p. 1).

Research into the field of syntax processing and aptitude generally observes great individual differences for the acquisition of syntactic knowledge (see for example Nauchi & Sakai, 2009; Hulstijn, 2005; Pakulak & Neville, 2010). Generally, ERP-[1] (see Pakulak & Neville, 2010; Tanner, Inoue & Osterhout, 2014) as well as FMRI-based studies (see Golestani et al., 2006; or Nauchi & Sakai, 2009) observed differences in the brain organisation between high and low proficiency groups tested through syntax processing tasks. The studies indicate activation in the left inferior frontal gyrus (Nauchi & Sakai 2009, p. 2626; Indefrey et al., 2004; Golestani et al., 2006, p. 1029), while some expressively emphasise the activation of the Broca's area situated in the inferior frontal gyrus (see for example Golestani et al., 2006, p. 1038; or Flöel et al., 2009, p. 1979). Some studies furthermore indicate a significant variability in white matter integrity around the Broca's between high and low proficiency groups in terms of grammatical aptitude (see Flöel et al., 2009, p. 1979). In essence, it is widely agreed upon that there is a significant correlation of syntactic ability and increased brain activity and differing brain structures.

[1] Event-related potential, which refers to the brain response as direct result of a stimulus. Regularly, EEGs are used to measure ERP.

1.5 Studies on Musicality and Grammatical Aptitude

As aforementioned, studies examining the relation of musicality and grammatical aptitude are relatively scarce. Maess and Koelsch have conducted various studies and experiments attempting to establish correlations between the processing of musical and language syntax, including an MEG study in 2001, suggesting that complex-rule based information is likely to be processed in the "Broca's area and its right-hemisphere homologue" for language and music: "from a functional-neuroanatomical view [there is] a strong relationship between the processing of language and music" (Maess & Koelsch, 2001, p. 543). Further ERP studies observed similar activations in the Broca's area (Koelsch 2005, p. 209; and Koelsch et al., 2005) while a study from 2011 remained inconclusive on the relation of music and language in a simultaneous processing setting (Maidhof & Koelsch, 2011). Kunert et al. conducted an fMRI study further suggesting interaction of music and language processing in the Broca's area (2015, p. 11). The authors however point out that this may just be limited to the processing of violations (2015, p. 12). In contrast to these studies, Slevc and Miyake observed no correlation between language aptitude and syntactical nor lexical knowledge in their study (2006, p. 679). Most recently, Gordon, Jacobs, Schuele and McAuley (2015) as well as Kalcheva and Fonseca-Mora (2017) observed strong links between music and grammar in their studies. Kalcheva and Fonseca-Mora examined two groups (singers with professional training and non-musicians) of adult Spanish learners of English and conclude: "our study points to a beneficial influence of singing on grammar achievement as part of the foreign language learning process in adults" (2017, p. 391). These most recent experiments further support the importance of exploring possible correlations of grammatical aptitude and musical skills.

1.6 Hypothesis

As demonstrated in the literature review, research on brain structures and activation of certain brain regions while processing musical and grammatical input indicates numerous similarities. Moreover, additional similarities could be expected through the processes of de- and encoding of sequences in music and the processing of novel grammar.

The underlying rationale of this study aims to consider the abilities intrinsic to musicality which are fostered and enhanced through active participation in a musical setting such as an ensemble, orchestra or band music, which should strongly influence pattern detection, retention and application.

The general, overarching hypothesis enquires about the overall correlation of musicality and grammatical aptitude, thus H_1 constitutes itself as follows:

H_1: There is a significant positive correlation between musicality scores and the grammatical aptitude test.

H_2 on the other hand specifically aims to observe correlations of grammatical aptitude and musical expertise and training, hence:

H_2: Musicians achieve a significantly higher score on grammatical aptitude tests than non-musicians.

Finally, the study is interested in the specific sub-skills of musicality and which of these relate to grammatical aptitude. The musicality test encompasses 4 subcategories: tuning, melody, accent and tempo. Melody and accent requires the participant to process, and retain rhythmic and melodic patterns, which is hypothesised to be closely linked to the encoding and decoding of grammatical syntax. Thus, the final hypothesis aims to observe the following correlation:

H_3: The musicality subtests that test pattern recognition and retention (Melody and Accent) correlate stronger with the grammatical aptitude test than the phonetic recognition tests (Tuning and Tempo)

2 Methodology

2.1 Participants

The sample of the study consisted of 25 participants, 14 male (56%) and 11 female (44%). The questionnaire divided the group into 10 musicians (40%) and 15 non-musicians (60%) based on musical experience in terms of training and active participation in a musical setting. However, it has to be noted that despite considerable efforts, the musician group sees a skewed sample size in regard of gender distribution: of the 10 musicians, only 2 females (20%) could be recruited for the study. In terms of age, the group can be regarded as rather homogenous with a mean of 30 years ($SD = 4.3$), the youngest being 23 and the oldest 35 years old.

Regarding education, the sample group again displays a considerable level of homogeneity, as 9 of the 25 participants (36%) are currently enrolled in university programs, whereas 16 (64%) already obtained at least one university degree.

All participants reported to be German native speakers and, based on the self-assessment of the questionnaire, spoke English at least on B2-level.[2] All participants reported to be proficient at least at one more language, ranging from A2 to C2 level.

[2] The Common European Framework of Reference has been used as basis, for more details on the respective levels of proficiency see: Council of Europe 2011. The Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge: CUP.

2.2 Instruments

To investigate the hypotheses, two tests were administered to assess musicality and grammatical aptitude. First of all, the Mini-Proms test (available online[3]) devised by Strauß et al. (2015) was used to determine overall musical abilities of the participants. The test assesses overall musicality through four subtests: *melody*, *accent*, *tuning* and *tempo*. Each section differed on maximum points achievable, *melody* and *accent* ranging from 0 to a maximum of 10 points, *tuning* and *tempo* from 0 to 8 points, resulting in a total of 36 points overall.

The *melody* section asks the test taker to compare harmonic sequences and patterns. The following section, *tuning*, requires the participant to determine if consecutive tones are the same or different in regard to pitch. Thirdly, *tempo* similarly requires the test taker to compare the relative tempo of two sound sequences with a monotone rhythm without any accents. And, finally, the *accent* section requires the retention and detection of rhythmic patterns over a sequence of 5–12 beats. According to Strauß et al., the test shows strong reliability and consistency throughout the entire procedure, being able to reliably predict musical experience and training (see Strauß et al., 2015).

Secondly, to assess grammatical aptitude, this study included section B of the Oxford Classics Language Aptitude Test (2013), a subtest of the Oxford University Classics Admissions Test. This specific sub-section features an artificial language, *Fub*, and requires the participants to detect, recognise and retain grammatical patterns and subsequently apply these rules in translation tasks ranging from short phrases to longer sentences featuring compound clauses. As the sample group had a diverse language background aside the shared L1 and English (including Hungarian, Spanish, Turkish, Croatian, Chinese, Dutch, French, or Italian) the artificial language allowed to eliminate any bias in relation to the language background. Moreover, the English skill level used in the prompt and the tasks themselves did not exceed B2 at any point, thus eliminating any further advantage based on the respective skill levels of English of the participants. Finally, the tasks closely follow the aforementioned construct of Purpura (2004, 91) eliciting grammatical knowledge pertaining to syntactical and morphological form. It has to be noted that scoring was done by the author of the study, attempting to eliminate any intra-rater bias through concealing the names of the test-takers and evaluating every test twice.[4] The test section provided the scores for each item independently, the maximum score

[3] Universität Innsbruck 2017. "Mini-Proms". https://www.uibk.ac.at/psychologie/fachbereiche/pdd/personality_assessment/proms/take-the-test/mini-proms/. (25 Jul. 2017).

[4] It has to be noted however, that the Oxford Classics Language Aptitude Test is not validated and generally not used for language aptitude testing. Moreover, while a solution sheet was provided, the scoring of the tests is still subject of personal assessment and consideration. Thus, intra- and inter-rater reliability may be considered average. Despite these reliability issues, construct and content validity and the relative authenticity of the language processing situations do outweigh, in the opinion of the author, the downsides of the scoring.

possible amounted to 50 points. Each item consisted of translation tasks, either requiring the participant to translate the artificial language into English or vice versa. The following example aims to illustrate the task:

huufls hohub red	The teachers taught the pupils.
hiip pik hohub	A teacher provided homework.
rored pik daawl	The pupils had homework.
rored pik liikl	The pupils liked the homework.
tok daaw rored	A pupil had a dog.
totok liiks red	The dog liked the pupils.
tok rored huuf	The pupil taught the dog.
paat pik totok	The dog ate the homework.

Give the meaning of:
liikl hohub tok
[3] (Oxford Classics Language Aptitude Test, 2013, Section B(a)).

Thus, each section commenced with a set of examples providing enough information to deduce morphological and syntactical patterns and to apply them in the consecutive translation tasks. The maximum points per task is given in the squared bracket, with each separate word (‘liikl‘, ‘hohub‘, and ‘tok‘) amounting to 1 point, half for correct vocabulary, and half for correct form (‘tok‘in the example representing the vocabulary ‘dog‘and the grammatical function of object of the sentence).

2.3 Procedure

A preliminary basic questionnaire was compiled aiming to assess the musical background of the participants. Aside from general questions determining age, gender, nationality, language background and language proficiency, the main corpus of the questionnaire elicited the amount and duration of musical training, as well as the participation in organised musical environments, such as band practice, orchestra or ensembles. In line with the proposed hypothesis, to be qualified as a musician, the participants had to attend a minimum amount of musical training (2 years in the last 5 years, or 5 years overall) in addition to regular participation in an orchestra or similar musical activity (again 2 years in the last 5, or 5 overall). Due to the small sample size of musicians, no distinction was made between instrumentalists, multi-instrumentalists, or vocalists.

Subsequently, the participants received an online link to test for musicality (details below in the test section) and, finally, an email with an editable pdf file including the grammar aptitude test. The participants were asked to send a screenshot of the final results page of the musicality test (including the detailed sub-scores) together with the filled-in grammar test back for assessment. The overall length of the two tests combined amounted to roughly 1 h and 15 min.

3 Results

A One-Sample Kolmogorov-Smirnov Test was conducted to test for a normal distribution of the musicality and grammar test. Results suggest a normal distribution for both tests (p = 0.20 for the Musicality Test and p = 0.55 for the Grammar Test), which was confirmed visually with QQ-plots and histograms.

To test for H_1, the correlation of musicality and grammatical aptitude, a Pearson-Correlation Test was conducted. Results show a strong, significant correlation ($r = 0.688$, $p < 0.001$) between the results of the musicality test scores and the grammatical test scores. This finding is supported visually by a scatterplot depicting the correlation between the musicality test scores and the grammar test ($r^2 = 0.474$, see Fig. 1). Hence, the H1 can be accepted.

To test for H2, T-Tests for independent samples for musicians and non-musicians was conducted with respect to both the musicality and grammar tests. Considering the musicality test, musicians scored on average 25.95 points ($SD = 5.1$) while non-musicians reached an average of 19.73 points (SD = 2.4) on a 36 points scale. Levene‘s Test for Equality of Variances between musicians and non-musicians for the musicality test shows that equal variances between the two groups can be assumed ($p = 0.57$). The T-test for equality of means showed a significant difference between the two groups ($t(23) = -3.58$; $p = 0.002$,).

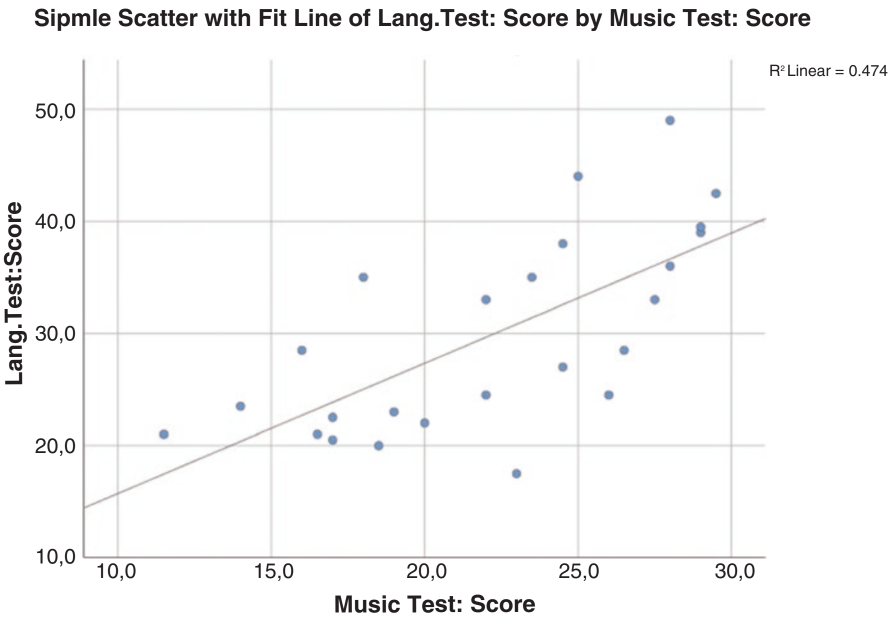


Fig. 1 Scatterplot of musicality test and grammar test

Similarly, the results of the grammatical aptitude test differed significantly between musicians (M = 36.6, SD = 6.76) and non-musicians (*M* = 25.47, *SD* = 6.76). Levene's Test revealed homogeneity of variance ($p = 0.976$) with the T-Test again showing a significant difference between musicians and non-musicians ($t(23) = -3.939$; $p = 0.001$). Hence, the results display a significant difference between the results on the musicality test as well as on the grammar test between the two groups of participants. Thus, H2 can be accepted.

Finally, H3 aimed to examine correlations between the subtests *melody* and *accent* of the musicality test in contrast with *tuning* and *tempo* and the grammar test. Pearson-Correlations of each sub-skill and the grammar test results display the expected divergences between the pattern detection and retention subsets of *accent* and *melody* compared to the phonetic perception tests, *tuning* and *tempo*. The subtest *accent* displayed the strongest, significant correlation ($r = 0.696$, $p < 0.001$) with the grammar test results, followed by *melody* ($r = 0.623$, $p = 0.001$), *tempo* ($r = 0.502$, $p = 0.01$) and *tuning* ($r = 0.476$, $p = 0.016$). Hence, while all four subskills correlate positively with the grammar test results, *melody* and *accent* clearly show a significantly stronger overall correlation. Hence, H3 can be accepted.

In order to evaluate possible factors interfering with the results, more tests were conducted. An independent samples T-Test comparing the results of males and females on the musicality and grammar test showed insignificant results regarding the musicality test ($t(23) = 1.124$; $p = 0.273$; male $M = 23.25$, $SD = 5.6$; female $M = 20.91$, $SD = 4.5$). In contrast, the results of the grammar test show significant differences between gender groups ($t(23) = 2.209$; $p = 0.037$; male $M = 33.11$, $SD = 9.4$; female $M = 25.86$, $SD = 6.2$). However, as elaborated on above, the sample group did not consist of matching numbers of females and males, especially concerning the musician group, which is most likely the reason for the differing results.

Regarding age and the musicality test, the Pearson test showed a significant positive correlation ($r = 0.416$, $p = 0.039$) while the relation of age and the language test was not significant ($r = 0.297$, $p = 0.149$). As Fig. 2 demonstrates however, generally musicians in the sample size were older than their non-musician counterparts, which again is very likely the reason for the observed differences.

Finally, in terms of education, an independent samples t-test demonstrated that there is no difference between the two groups present (university degree and Matura (i.e. Austrian A-levels)) for both the musicality test ($t(23) = -0.247$; $p = 787$) and language test ($t(23) = -0.316$; $p = 0.755$).

Overall, the results strongly indicate correlations between musical expertise and grammatical aptitude. Additionally, H3 suggests that musical abilities that are related to pattern analysis, retention and application also correlate more strongly with the results of the grammatical aptitude test in comparison to the detection of phonetic characteristics in terms of pitch and tempo. Naturally, the observed results only indicate correlations and cannot attest for any causality.

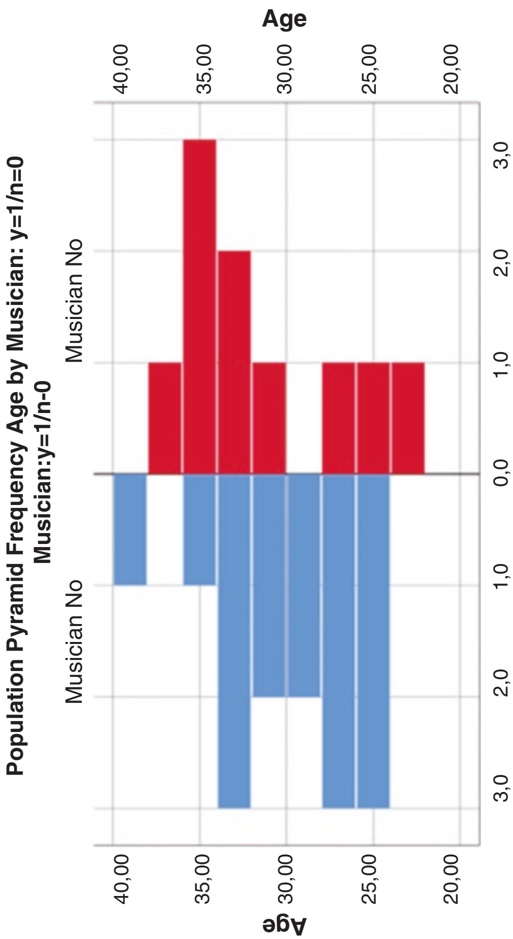


Fig. 2 Population pyramid highlighting age distribution among musicians and non-musicians

4 Discussion

The previously listed research offers a plethora of explanations for these observed correlations. As discussed above, studies on grammatical aptitude and musicality regularly highlight similar brain regions involved in the processing of music and language. Maess et al., having observed activations in the Broca‘s area while processing musical chords, indicate that

> complex rule-based information is processed in these areas with considerably less domain-specificity than previously believed. This might suggest that these areas process syntax, that is, complex rule-based information, in a domain other than language (2001, p. 543)

Similarly, numerous studies emphasise similarities in the nature of processing music and syntax information in language through the shared requirement of pattern detection. Kunert et al. ascertain that "[i]nstrumental music and language are both syntactic systems, employing complex, hierarchically-structured sequences built using implicit structural norms" (2015, p. 1), while Brown et al. regard music and speech both as combinatorial systems "in which larger structures are generated hierarchically from a pool of smaller, more unitary components" (2006, p. 2791). Also Flöel et al. emphasise these factors, indicating that the acquisition of grammar implies the "extraction of rule-based information" (2009, p. 1974). This again can be related to the processing of music, as Kraus and Chandrasekaran point out:

> Active engagement with music improves the ability to rapidly detect, sequence and encode sound patterns. Improved pattern detection enables the cortex to selectively enhance predictable features of the auditory signal at the level of the auditory brainstem (2010, p. 600)

These observations generally highlight that musicality is not limited to the processing of sounds alone. Aside from the auditory skills involved, active musical engagement naturally also relates to motor skills, but also to the processing of rules and patterns (see Kalcheva & Fonseca-Mora, 2017, p. 391). This may also explain the underlying rationale behind H_3, the assumed positive correlation of the pattern-based musicality sub-tests, melody and rhythm. These two sub-skills are obviously core characteristics of musical interaction. Hence, the results of the present study indicating a relation between musical expertise through practice and training in an interactional setting and grammatical aptitude could in fact be based on similar processes of pattern detection and extraction and application of deduced rules.

Furthermore, Fonseca-Mora, Toscano-Fuentes and Wermke emphasise that, after all, "language acquisition depends on interaction" (2011, 101). It can be assumed that increased interaction and focus through practice may increase the success of language acquisition. Similarly, this study defined the prerequisite to be qualified as musician as having participated in an organised musical setting which naturally involves interaction as well. The interactional nature and the processes involved in musical interaction per se are lacking conclusive research, thus any relation based on similar interactional settings can only be hypothesised. However, to add to the interactional nature of language and music, simple personal experience can attest for the complex process of interacting with others through music. Clearly, musical communication predominantly requires the processing of auditory cues and input, but it also stresses rhythmic and melodic pattern detection to anticipate and comprehend structures to enable the musicians to interact successfully. From this point of view, the processing of language and music can be regarded as a similar process of inferencing, de- and encoding of information followed by the application of these complex rule-based in the generation of another sentence or the next part of a sequence of sounds and rhythmic patterns. Chobert and Besson support this notion by emphasising that "musical practice requires sustained attention control and memory" (2013, p. 931) which in turn may positively affect language processing as well. Finally, the aforementioned bootstrapping theory may yet add another line of reasoning for the correlation of musicality and grammatical aptitude. Soderstrom et al. indicate that infants may deduce syntactic boundaries of strings of language "even before lexical knowledge is available" through a natural sensitivity to prosodic markers of "syntactic units smaller than the clause" (2003, p. 249). Mazuka (2007)support these findings in their own experiments, adding that the rhythmic organisation of a language "provides the child learner with a means of segmenting the speech stream into linguistically significant units (2007, p. 1). However, it has to be noted that the bootstrapping theory is highly disputed. Fernald and McRoberts criticise the absence of direct evidence and the problematic inconsistency of acoustic cues (1996, p. 365).

Naturally, numerous other factors can be considered to have caused the observed results and correlations. First of all the present study does contain certain incalculable factors due to administrative difficulties. Primarily, the small sample size may negatively impact the overall validity. Especially considering that the participants took the tests in an uncontrolled environment, the study cannot attest for factors such as external help, extension of the time limits or participant-related reliability factors such as fatigue.

Furthermore, as elaborated on above, the test itself displays certain issues concerning rater-reliability. Additionally, the Oxford language aptitude test did not undergo a validation process, thus construct and content validity cannot be accounted for. However, it has to be noted that the test is used in this form as admissions test since a considerable amount of time and the tasks themselves strongly suggest indicate construct validity in relation to the framework suggested by Purpura (2004, p. 91).

Secondly, as Strait and Kraus quite poignantly observe, the general problem with comparative studies is the problem of other interfering factors that can hardly ever be attested for, such as general intelligence, socio-economic background or learning methods and strategies (2011, p. 133). Thus, it is entirely possible that the groups examined in the present study may also differ significantly on any these factors, as they are all regularly linked to aptitude and language processing as well. The design of the study made it impossible to exclude these possible intervening factors, however, larger scale studies could include intelligence, working memory and further additional tests to better control these factors. Despite these limitations, the sample group at least showed general homogeneity in respect of educational background, and common L1, similar levels of L2.

Especially the relation of working memory and language aptitude demands further mentioning. While research investigating this relationship is still limited, quite a few studies show indeed a positive correlation (see for example Yalçın, Sevdeğer & Erçetin, 2016). Yalçın, Sevdeğer and Erçetin furthermore observed a strong correlation between working memory and grammatical inferencing (2016, p. 144). Moreover, a comprehensive meta-analysis of 79 studies in 2013 strongly supports a positive relationship between working memory and L2 proficiency (see Link et al. 2014). The grammatical aptitude test does represent itself as a rather fitting example for working memory and aptitude relations due to the process of the test involving retention of phrases under time constraints.

Furthermore, another factor that may have influenced the results on the grammar test could be the motivational aspect of the study. The structure of the testing sequence allowed the participants to see their results on the musicality test (together with a short summary evaluating their scores) before taking the grammar test. Dörnyei and Skehan specifically emphasise that motivational aspects are a very strong predictor of learning success (2003: 589). Thus, the experience and the results of the musicality test may have influenced motivation and performance on the following task.

Finally, another factor possibly influencing results, is reading ability, which may in fact be related to musicality as a growing number of studies suggest. Overy et al. observed correlations between the detection of musical timing and tempo perception and dyslexia, concluding that reading impairment may be remedied through musical training (2003, p. 34). Similarly, Strait, Hornickel and Kraus support the relation of musical aptitude and general reading ability, concluding that their data acquired through a small scale empirical study among school children indicate "common brain mechanisms underlying reading and music abilities that relate to how the nervous system responds to regularities in auditory input" (2011, p. 1). Another recent small scale study adds to these observed correlations: Bekius, Cope and Grube conclude that their findings confirm the "relevance of auditory regularity processing in reading skill" (2016, p. 8). As emphasised above, every grammar or vocabulary task has to be embedded in a skill such as reading, speaking, listening or writing. Thus, while it seems impossible or impractical to attempt to eliminate this factor, a reading ability test could be administered to control for this factor as well.

The relatively long list of limitations and various other factors possibly influencing test results only highlights the complexity of language aptitude and the concept of musicality. Despite all these aforementioned factors, especially working memory, motivation and reading which may have contributed to the overall test results, the overall strong correlations of the presumed hypotheses relating musicality and grammatical syntax can not be disregarded.

5 Conclusion

The present study highlights the possible relation of musical expertise and grammatical aptitude. Despite various factors that could not be accounted for, the results definitely warrant further research on these correlations. The review of previous studies also highlights that the concepts of musicality, grammar, and aptitude in general require considerable research to better grasp and define these terms. The results of the study can be based on presumed similarities of language and music processing in regard of pattern de- and encoding. Clearly, musicality and active musical engagement encompass more than the processing of sounds. Especially the focus on patterns transmitted through rhythm, accentuation, and harmonic progression indicate the complex nature of musicality. Thus, the present study strongly suggests further research into these rather unrepresented aspects of musical ability and language processing. Naturally, the present study can only add a small indication of these possible relationships, especially considering the inherent limitations. Eventually, consecutive studies certainly have to account for the, frankly, numerous factors that may influence language aptitude while testing for correlations between musicality and language processing.

References

Bekius, A., Cope, T. E., & Grube, M. (2016). The beat to read: A cross-lingual link between rhythmic regularity perception and reading skill. *Frontiers in Human Neuroscience, 10*(425), 1–11. https://doi.org/10.3389/fnhum.2016.00425

Bermudez, P., Lerch, J., Evans, A., & Zatorre, R. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex, 19*(7), 1583–1596. https://doi.org/10.1093/cercor/bhn196

Brown, D., & Abeywickrama, P. (2010). *Language assessment: Principles and classroom practices*. New York: Pearson Longman.

Brown, S., Martinez, M., & Parsons, L. (2006). Music and language side by side in the brain: A PET study of the generation of melodies and sentences. *European Journal of Neuroscience, 23*(10), 2791–2803. https://doi.org/10.1111/j.1460-9568.2006.04785.x

Carroll, J. (1962). The prediction of success in intensive foreign language training. In Glaser R. (Ed). (2009). *Training, research, and education*. Pittsburgh, PA: University of Pittsburgh Press.

Carroll, J. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *International Journal of PsychoLinguistics, 2*, 5–14.

Carroll, J., & Sapon, S. (1959). *The modern languages aptitude test*. San Antonio, TX: Psychological Corporation.

Chobert, J., & Besson, M. (2013). Musical expertise and second language learning. *Brain Sciences, 3*(2), 923–940. https://doi.org/10.3390/brainsci3020923

Christiner, M., & Reiterer, M. S. (2015). A Mozart is not a Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Human Neuroscience, 9*(482), 1–8. https://doi.org/10.3389/fnhum.2015.00482

Council of Europe. (2011). *The common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, MA: Cambridge University Press.
Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In Doughty, C. J., & Long, M. H. (Eds.). (2003). *The handbook of second language acquisition*. Malden, MA: Blackwell Publishing.
Fernald, A., & Mcroberts, G. (1996). Prosodic bootstrapping: A critical analysis of the argument and the evidence. In Morgan, J., & Demuth K. (Eds.). (1996) *Signal to syntax: Bootstrapping from speech to grammar in early acquistion*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
Flöel, A., de Vries, M., Scholz, J., Breitenstein, C., & Johansen-Berg, H. (2009). White matter integrity in the vicinity of Broca's area predicts grammar learning success. *NeuroImage, 47*(4), 1974–1981. https://doi.org/10.1016/j.neuroimage.2009.05.046
Fonseca-Mora, C., Toscano-Fuentes, C., & Wermke, K. (2011). Melodies that help: The relation between language aptitude and musical intelligence. *International Journal of English Studies, 22*(1), 101–118.
Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience, 23*(27), 9240–9245. https://doi.org/10.1016/S1053-8119
Golestani, N., Alario, X., Meriaux, S., Le Behian, D., Dehaene, S., & Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia, 44*(7), 1029–1040. https://doi.org/10.1016/j.neuropsychologia.2005.11.009
Gordon, R., Jacobs, M., Schuele, M., & Mcauley, D. (2015). Perspectives on the rhythm-grammar link and its implications for typical and atypical language development. *Annals of the New York Academy of Sciences, 1337*, 16–25.
Grigorenko, E., Sternberg, R., & Ehrman, M. (2000). A theory- based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal, 84*(3), 390–405.
Honing, H., ten Cate, C., Peretz, I., & Trehub, S. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transaction B, 370*(1664), 1–8. https://doi.org/10.1098/rstb.2014.0088
Hughes, A. (2003). *Testing for language teachers*. Cambridge, MA: Cambridge University Press.
Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition, 27*(02), 129–140. https://doi.org/10.1017/S0272263105050084
Indefrey, P., Hellwig, F., Herzog, H., Seitz, R., & Hagoort, P. (2004). Neural responses to the production and comprehension of syntax in identical utterances. *Brain and Language, 89*(2), 312–319. https://doi.org/10.1016/S0093-934X(03)00352-3
James, C., Oechslin, M., van de Ville, D., Hauert, C., Descloux, C., & Lazeyras, F. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Structure and Function, 219*(1), 353–366. https://doi.org/10.1007/s00429-013-0504-z
Jilka, M. (2009). Talent and proficiency in language. In Dogil G., & Reiterer M. S. (Eds.). (2009). *Language talent and brain activity*. Berlin: Mouton De Gruyter.
Kalcheva, I., & Fonseca-Mora, M. C. (2017). *Musical experience in singing and grammar achievement in English language acquisition*. Foreign languages and the contemporary higher education conference proceedings VIII, International Scientific Conference, 23–25 June 2017, Varna, Bulgaria, 391–397.
Kepinska, O., de Rover, M., Caspers, J., & Schiller, N. (2016). On neural correlates of individual differences in novel grammar learning: An fMRI study. Neuropsychologia, 98, 156–168. doi:https://doi.org/10.1016/j.neuropsychologia.2016.06.014.
Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology, 15*, 207–212. https://doi.org/10.1007/978-3-211-75121-3_9

Koelsch, S., Gunter, T., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience, 17*(10), 1–13. https://doi.org/10.1162/089892905774597290
Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience, 11*(8), 599–605. https://doi.org/10.1038/nrn2882
Kunert, R., Willems, R., Casasanto, D., Patel, A., & Hagoort, P. (2015). Music and language syntax interact in Broca's area: An fMRI study. *PLoS One, 10*(11), 1–16. https://doi.org/10.1371/journal.pone.0141069
Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics, 36*(3), 385–408. https://doi.org/10.1093/applin/amu054
Link, J., Osthus, P., Koeth, J., & Bunting, M. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review, 21*(4), 861–883. https://doi.org/10.3758/s13423-013-0565-2
Maess, B., Koelsch, S., Gunter, T., & Friederici, A. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4*(5), 540–545. https://doi.org/10.1038/87502
Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience, 23*(9), 2252–2267. https://doi.org/10.1162/jocn.2010.2154
Marcus, G. (2012). Musicality: Instinct or acquired skill? *Topics in Cognitive Science, 4*(4), 498–512. https://doi.org/10.1111/j.1756-8765.2012.01220.x
Marie, C., Delogu, F., Lampis, G., Olivetti, B. M., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in mandarin Chinese. *Journal of Cognitive Neuroscience, 23*(10), 1–15. https://doi.org/10.1162/jocn.2010.21585
Marques, C., Moreno, S., Luís, C. S., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience, 19*(9), 1453–1463. https://doi.org/10.1162/jocn.2007.19.9.1453
Mazuka, R. (2007). The rhythm-based prosodic bootstrapping hypothesis of early language acquisition: Does it work for learning for all languages? *Gengo Kenkyú, 132*, 1–11.
Meara, P. (2005). *LLama language aptitude tests*. Swansea: Lognostics.
Milovanova, R., Huotilainenc, M., Välimäkid, V., Esquefd, P., & Tervaniemic, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research, 1194*, 81–89. https://doi.org/10.1016/j.brainres.2007.11.042
Nauchi, A., & Sakai, K. (2009). Greater leftward lateralization of the inferior frontal gyrus in second language learners with higher syntactic abilities. *Human Brain Mapping, 30*(11), 3625–3635. https://doi.org/10.1002/hbm.20790
Overy, K., Nicolson, R., Fawcett, A., & Clarke, E. (2003). Dyslexia and music: Measuring musical timing skills. *Dyslexia, 9*(1), 18–36. https://doi.org/10.1002/dys.233
Oxford Language Aptitude Test (and Solution Form). (2013). Oxford University. Retrieved from http://www.classics.ox.ac.uk/tests.html?file=tl_files/Downloads/CLAT%20Tests/CAT%202013.pdf. http://www.classics.ox.ac.uk/tests.html?file=tl_files/Downloads/CLAT%20Tests/CLAT%202013%20solutions.pdf.
Pakulak, E., & Neville, H. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience, 22*(12), 2728–2744. https://doi.org/10.1162/jocn.2009.21393
Pastuszek-Lipinska, B. (2004). An overview of a research project and preliminary results of two experiments on perception and production of foreign language sounds by musicians and nonmusicians. *TMH-QPSR, 46*(1), 61–74.
Patel, A., Gibson, E., Ratner, J., Besson, M., & Holocomb, P. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*(6), 717–733. https://doi.org/10.1162/089892998563121

Pei, Z., Wu, Y., Xiang, X., & Qian, H. (2016). The effects of musical aptitude and musical training on phonological production in foreign languages. *English Language Teaching, 9*(6), 19–29. https://doi.org/10.5539/elt.v9n6p19

Pimsleur, P. (1966). *The Pimsleur language aptitude battery*. New York: Harcourt, Brace, Jovanovic.

Purpura, J. (2004). *Assessing grammar*. Cambridge, MA: Cambridge University Press.

Schoen-Nazzaro, M. (1978). Plato and Aristotle on the ends of music. *Laval théologique et philosophique, 34*(3), 261.

Slevc, R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science, 17*(8), 675–681. https://doi.org/10.1111/j.1467-9280.2006.01765.x

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage, 17*(3), 1613–1622. https://doi.org/10.1006/nimg.2002.1288

Soderstrom, M., Seidl, A., Kemler, N., Deborah, G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from Prelinguistic infants. *Journal of Memory and Language, 49*(2), 249–267. https://doi.org/10.1016/S0749-596X(03)00024-X

Strait, D., Hornickel, J., & Kraus, N. (2011). Subcortical processing of speech regularities underlies reading and music aptitude in children. *Behavioral and Brain Functions, 7*(44), 1–11. https://doi.org/10.1186/1744-9081-7-44

Strait, D., & Kraus, N. (2011). Playing music for a smarter ear: Cognitive, perceptual and neurobiological evidence. *Music Perception, 29*(2), 133–146. https://doi.org/10.1525/MP.2011.29.2.133

Strauß, H., Shakeshaft, N., Plomin, R., & Zentner, M. (2015). Introducing the mini-proms: A short and objective measure of musical ability. International convention of psychological science (ICPS), 12–14 March 2015, Amsterdam. Retrieved from https://www.uibk.ac.at/psychologie/fachbereiche/pdd/personality_assessment/research/download/minipromsposter_2015.pdf

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition, 17*(2), 277–293. https://doi.org/10.1017/S1366728913000370

Universität Innsbruck. (2017). *Mini-Proms*. Retrieved from https://www.uibk.ac.at/psychologie/fachbereiche/pdd/personality_assessment/proms/take-the-test/mini-proms/

Volz, M. (2005). Improvisation begins with exploration. *Music Educators Journal, 92*(1), 50–53. https://doi.org/10.2307/3400227

Yalçın, Ş., Sevdeğer, Ç., & Erçetin, G. (2016). The relationship between aptitude and working memory: An instructed SLA context. *Language Awareness, 25*(1), 1–15. https://doi.org/10.1080/09658416.2015.1122026

Language Aptitude and Gender

Cornelia Habl

Abstract In recent years, research has shown that the neurocognition of language, e.g. how language is learned, stored and retrieved, differs between men and women. To a certain extent, these findings could be linked to differences in brain structure and function, especially the gender specific connectivity of the brain and convergent activation patterns during speech perception. Women were found to show greater interhemispheric activity than men in language related tasks, pointing towards a stronger emotional involvement in declarative memory retrieval. In men, the same processes seem to be performed by only one hemisphere which is connected to the procedural memory system, specialised for rules and sequences. Considering the bilateral hemispheric activation in women in language related tasks, this study firstly hypothesised that they would outperform men in vocabulary memory tasks. Due to the assumption that men rely more on the procedural memory system for memorising lexical information, the second hypothesis examined whether men perform better in grammar learning tasks than women. Using the LLAMA B and the LLAMA F test, participants were required to remember words or grammatical rules for artificial languages respectively. Even though the results show no significant differences, they yield interesting points for discussion about possibly undesirable links between the two tests. Furthermore, interviews with the participants after testing in relation to their test scores showed fascinating links between testing success and emotional involvement with the stimuli.

1 Introduction

Scientists found gender differences regarding language learning, storing and retrieving, meaning that the neurocognition of language differs between males and females (Ullman et al., 2008). One aspect of these gender differences observed in language

C. Habl (✉)
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Austria
e-mail: cornelia.habl@gmail.com

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_12

related tasks is of an anatomical and structural nature. Modern technology such as MRI and CT scans reveals several areas of the brain that differ in size between the sexes, and that different structures are used when solving language related tasks (Kansaku, Yamaura, & Kitazawa, 2000). In relation to brain structure, several hypotheses are also concerned with the evolutionary aspect of why human brains are structured the way they are, especially with regard to the *Homo sapiens'* Neolithic hunter/gatherer lifestyle and the gender specific work division that possibly went along with it.

Based on these anatomical and physiological factors influencing gender differences in language aptitude, a plethora of studies have been conducted to assess where these differences lie exactly and how they have come to pass from an evolutionary perspective. In this paper, I will present a small study on this topic and discuss the findings under the aforementioned aspects pertaining to gender differences in language aptitude.

Language learning aptitude is often referred to as the "prediction of how well, relative to other individuals, an individual can learn a foreign language in a given amount of time and under given conditions" (Stansfield, 1989, p. 438). Different aspects influencing this ability will be discussed in this section. Before, I would like to mention that my use of the term *gender* instead of *sex* throughout this report is owed to political correctness, and as all the participants' sex corresponded with their gender identity in this study, terminology does not obscure the results in any way. However, when quoting and reporting about other scientists' studies, I will use the term researchers were working with in their studies, as it might have made a difference in their choice of participants.

1.1 Anatomical and Physiological Differences in Male/Female Brains

While no difference between males and females in general intelligence can be reported, research has shown that often one gender performs certain cognitive tasks better than the other (Baron-Cohen, Knickmeyer, & Belmonte, 2005). Due to these findings on a population level, biological reasons for these differences seem plausible. Sex-related differences are visible in morphometric brain imaging. Research showed that male brains are 9% larger than female brains, and that this extra material is mostly white matter. This increased brain volume can mostly be explained by larger body size in males, and it is also connected to a decreased interhemispheric connectivity and a tighter packing of neurons, which results in increased local connectivity. The part of the brain which connects both hemispheres of the brain, the corpus callosum, is verifiably smaller in males than in females (Baron-Cohen et al., 2005). These structural differences led to several physiological studies where it was shown that females have greater bilateral language related activation than males (Baxter et al., 2003; Shaywitz et al., 1995). Consequently, a number of researchers have suggested that the male brain is more strongly lateralised than the female brain. When Kansaku et al. (2000) read a story to males and females in a functional MRI scanner, males showed a strong left lateralisation, while females showed bilateral activation of brain areas. Shaywitz et al. (1995) also reported stronger left lateralisation in males but restricts his findings to phonological processes only. Interestingly, McGlone (1977) reported that males with left hemisphere focal lesions were three times more likely to develop language deficits than females with left hemisphere lesions.

An activation of the left hemisphere in all language involved processes is to be expected since the two areas that are strongly connected to language production and processing, Wernicke's and Broca's area, are located in the temporal and frontal regions. While the former is classically viewed as being responsible for the comprehension and understanding of written and spoken language, the latter is associated with motoric aspects of speech production. Studies have shown that both areas are larger in females. They have 23% more volume in Broca's and 13% more volume in Wernicke's area (Schlaepfer et al., 1995).

Baron-Cohen et al. (2005) state that connectivity in the brain is linked to the ability to empathise. The fact that females show interhemispheric activity when confronted with language related tasks could imply an emotional involvement in declarative memory retrieval.

Additionally, sex differences exist in every lobe of the brain, including many 'cognitive' regions such as the hippocampus, amygdala and neocortex. Concerning the issue of lateralisation, the amygdala nuclei are interesting because of their role in emotional memory storage and modulation of memory strength in connection with stress hormones. They are not only significantly larger in males, but specifically the left hemisphere amygdala is more frequently used for visual, emotional memory in females, while males use the right hemisphere amygdala for the same processes. This hemispheric laterality of function parallels the aforementioned relationship of the amygdala with memory for emotional information (Cahill, 2006).

What is more, the frontal/basal ganglia circuits have been associated with the procedural memory system, which is specialised for rules and sequences, while the hippocampus and para-hippocampal regions and amygdala are connected to the declarative memory system and emotional memory. Following the above-mentioned studies, scientists believe that males rely more on the procedural memory system for memorising lexical information, while females use the declarative memory system for lexical knowledge storage (Ullman et al., 2008).

A psychological theory that ties in with the declarative/procedural memory system is the empathising-systemising (E-S) theory, according to which males are generally thought to be more systemising and females more empathising when the differences of the two dimensions are compared (Baron-Cohen et al., 2005). As Baron-Cohen et al. (2005) explain:

> Systemizing is the drive to analyze a system in terms of the rules that govern the system, in order to predict the behavior of the system. Empathizing is the drive to identify another's mental states and to respond to these with an appropriate emotion, in order to predict and to respond to the behavior of another person. (p. 820)

The systemising and empathising structures of our brains can be classified into categories and individual brain types (e.g. more systemising or empathising brains), according to the difference of the two dimensions along a continuum (Baron-Cohen et al., 2005).

Another physiological aspect that is connected to brain structure, and therefore language related task performance, is sex hormones like testosterone and oestrogen. A large amount of evidence indicates that oestrogens affect declarative memory and (para) hippocampal function, and that they can enhance performance on a variety of declarative memory tasks in women (Phillips & Sherwin, 1992; Sherwin, 1988). Most prominently, the hormonal influences from the menstrual cycle have been the subject of a multitude of studies, showing that both learning and memory processes are substantially influenced by sex hormones in combination with stress hormones. In humans, the stage of the menstrual cycle significantly influences performance on spatial and verbal tasks (Halpern & Tan, 2001), the degree of brain asymmetry involving cognitive tasks (Hausmann, 2005), and the responsiveness to addictive drugs like cocaine (Kaufman et al., 2001).

However, as several studies show, it is not only sex hormones that are responsible for physiological differences in male and female brains. A study on rats, a well-established human model, showed sex differences in cocaine-seeking behaviour and no effect of the oestrus state on these differences (Fuchs, Evans, Mehta, Case, & See, 2005).

1.2 Evolutionary Implications

The human body and therefore also the brain are the product of thousands of years of selection for structures and functions which are optimally adapted to humans' natural and social environment. It is therefore also the product of inter- and intra-sexual selection, meaning the competition for mates between and within the sexes (Cahill, 2006). Males and females have evolved different strategies for optimising mating success, although females tend to compete with other females more subtly, for instance via social cues. Ullman et al.'s (2008) reports on females' heightened recall of detailed information and non-verbal stimuli like faces, as well as Sherwin's (1988) observations on oestrogens influencing declarative memory, are in line with this theory.

It was also suggested that while males were able to solve conflicts with physical force due to their stronger physique, females were forced to rely on social manipulations to gain advantages in social groups, which goes hand in hand with superior declarative memory skills (Sabbatini, 1997).

Another popular theory about the evolution of gender specific skills involves the Neolithic hunter/gatherer lifestyle of our ancestors. During navigation, females rely more on landmark cues while males tend to rely more on geometric cues (Bever, 1992; Sandstrom, Kaufman, & Huettel, 1998; Saucier, Bowman, & Elias, 2003; Williams, Barnett, & Meck, 1990). In accordance with the respective brain structures, males are thought to be better at tasks that require spatial skills, depth reckoning and visuospatial processing for which they use frontal/basal ganglia circuits which have been associated with the procedural memory system specialised for rules and sequences (Ullman et al., 2008). Females' reliance on landmark cues ties in with their ability to internally verbalise these stimuli and their superior declarative memory skills, a system that may be particularly important for learning idiosyncratic information, specifically arbitrary relations (Lewin, Wolgers, & Herlitz, 2001; Saucier et al., 2003).

Ullman et al. (2008) suggest that the gender specific strategies in spatial navigation tasks can be linked to language storage and processing. "If females have superior declarative memory abilities as compared to males, we might expect females to be more likely to rely on stored complex forms, while men depend more on rule-based composition" (Ullman et al., 2008, p. 298). These assumptions have been the subject of extensive research.

1.3 Studies and Research on Language Aptitude and Gender

In several ontogenetic studies female children have been reported to perform better than their male contemporaries. Huttenlocher, Haight, Bryk, Seltzer, and Lyons (1991) conducted a study on 16–24 months old infants and reported that female infants have a larger vocabulary than males. In a study with 2–4 years old children, girls produced more complex utterances than boys (Horgan, 1975). Several studies support the view that these superior vocabulary abilities continue after this age, but that different strategies are used by males and females (Ullman et al., 2008).

Kimura (1999) reported that while females recall words in clusters and meaningful categories, males tend to recall words in the order they were presented in, which ties in with Ullman et al.'s (2008) claim that navigational and language learning skills are linked.

Specific focus has been put on so called episodic memory tasks, where subjects are asked to remember certain stimuli in a short period of time. Females outperformed males when asked to remember all kinds of verbal stimuli like words, digits or paragraph content (Kimura, 1999), but also nameable items like landmarks (Saucier et al., 2003) and real objects (Ullman et al., 2008). This verbal/spatial link does not, however, account for data pertaining to females' superior memory of object locations (Alexander, Packard, & Peterson, 2002), novel faces (Lewin et al., 2001) and complex abstract patterns (McGivern et al., 1997).

1.4 Hypotheses

The empirical study in this chapter is based on Ullman et al.'s (2008) assessment that females have superior declarative memory abilities and are therefore better at learning idiosyncratic information and arbitrary relations than males (also Lewin et al., 2001). Hence, it is assumed that females should be better at vocabulary learning than males. The first hypothesis is therefore that females perform better than males at learning unknown words and therefore score higher on the LLAMA B test.

According to Ullman et al. (2008), males have superior procedural memory abilities and should therefore be better at learning grammatical rules. Thus, the second hypothesis assumes that males are better at learning grammatical rules than females and score higher in the LLAMA F test.

2 Methods

2.1 Participants

I tested 18 people, 7 males and 11 females. The small sample size as well as the unequal gender distribution of the participants is owed to the limited time frame during which data collection had to occur and not to any methodological considerations. Table 1 below provides an in-depth description of the participants' backgrounds.

2.2 Instruments

For the empirical research, I used the downloadable versions of the LLAMA B and the LLAMA F test (Version 2.00, Llama Language Aptitude tests [Computer Software] (2015)), because these tests were the best fit for both my research questions and the prospective participants of my study. Firstly, I wanted to test vocabulary learning ability, for which the LLAMA B test became my method of choice due to the high level of comparability of test results. The LLAMA F test was chosen to test for grammatical rule learning ability due to the sequential nature of its test elements. Furthermore, both tests use artificial languages, making the first language of the participants inconsequential to the testing process (other than in the MLAT, for example).

Table 1 Details of male and female participants of the study

	Males	Females
N	7	11
Age	Range: 21–37, mean: 26.9y	Range: 21–29, mean: 25.8y
Nationality	Austrian (6), British (1)	Austrian (9), Greek (1), Luxembourg (1)
Degree	A-Levels (4), BA (1), MA (2)	A-Levels (2), BA (4), MA (5)
Nr of L2	Range: 1–3, mean: 2.0	Range: 1–4, mean: 2.8
Nr of countries	Range: 1–2, mean: 1.1	Range: 1–4, mean: 2.3
Nr of dialects	Range: 0–4, mean: 1.1	Range: 0–2, mean: 1.1

2.3 Procedures

The testing procedure required participants to firstly fill out a consent form and a short personal data sheet. They then received a brief tutorial on how to do the tests (LLAMA B was always administered first), but were not told what was actually tested. In the tests, participants were asked to remember artificial words and match them to pictures (LLAMA B) or infer grammatical rules of an artificial language (LLAMA F), all within a time limit of 2 min (LLAMA B) and 5 min (LLAMA F). After the memory or inference phase, participants had to click on the correct picture matching a word (LLAMA B) or chose between two sentences to describe a picture (LLAMA F). In both tests, the maximum score was 100 points.

Participants were left alone when the test was administered, but were encouraged to call for help if needed. All participants were tested in their private homes. Afterwards I asked about the participants' strategies for remembering the words and which test they thought was easier. Further statistical analysis was conducted with SPSS (Version 22).

3 Results

Concerning the mean test scores, females achieved higher scores than males in both the LLAMA B (M = 62.27, SD = 27.60) and the LLAMA F test (M = 67.27, SD = 21.02). Males performed better in the LLAMA B (M = 60.71, SD = 11.34) than in the LLAMA F test (M = 52.14, SD = 19.55). The test score distributions of both genders for the LLAMA B test show that females had both the highest and the lowest scores (15 and 95 points), while males had mostly average scores (40–70 points). Similarly, the test score distributions for the LLAMA F test show that females were responsible for the highest scores (3 participants scored between 90 and 100 points), while 6 out of 7 males scored between 40 and 80 points. A Kolmogorov-Smirnov Test revealed a normal distribution for all test scores for both genders (p = 0.149; p = 0.200).

The boxplot in Fig. 1 below displays the test scores (0–100) on the y-axis and gender on the x-axis. The boxes represent 90% of the scores, while the whiskers show the remaining 10% with the most extreme values at their end points. The bold line in the boxes represents the median scores of a test in a particular gender. The boxplot shows again that for the LLAMA B scores, females scored in a much wider range than males (90% of scores between 40 and 85 points versus 55–70 points). In contrast, the scores of the LLAMA F test showed an equal range in males (90% of scores between 40 and 70 points) and in females (90% of scores between 50 and 80 points). However, because the overlap between the two boxes is only partial, it points toward a trend for a difference between the scores of the two genders.

However, the t-test for independent samples did not reveal a significant difference between males (M = 52.14, SD = 19.548) and females (M = 67.27, SD = 21.019)

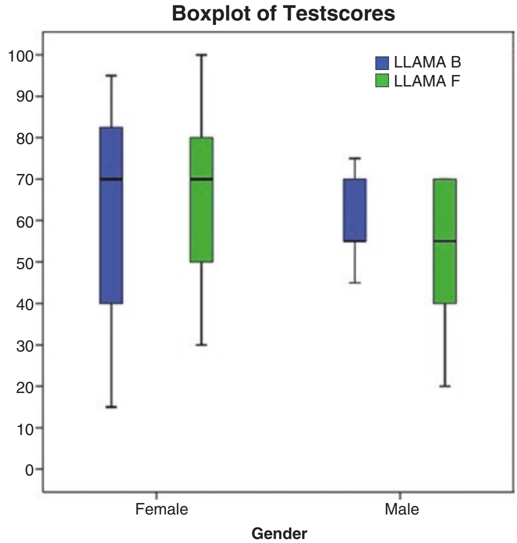


Fig. 1 Boxplot of test scores of LLAMA B and LLAMA F scores

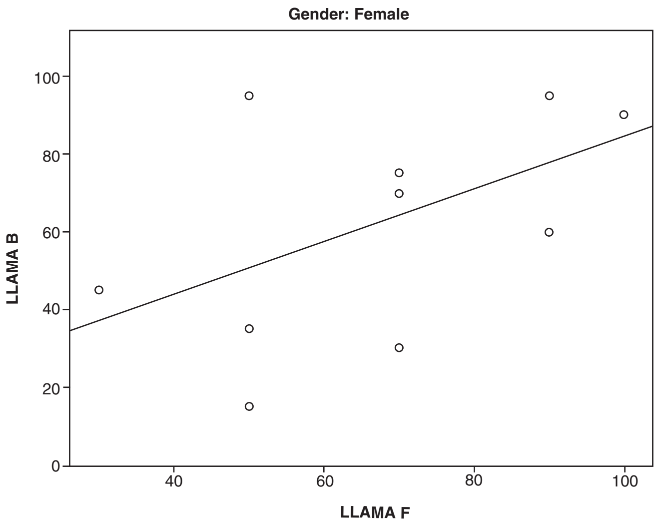


Fig. 2 Scatterplot of a positive correlation (Spearman-Rho) between LLAMA B and LLAMA F test results in females

for the LLAMA F; $t(16) = 1.528$, $p = 0.146$. Likewise, the results were insignificant between males ($M = 60.71$, $SD = 11.339$) and females ($M = 62.27$, $SD = 27.601$) for the LLAMA B test; $t(14.328) = 0.166$; $p = 0.870$.

An analysis of correlations (Spearman-Rho) between other variables in females showed a positive trend between the LLAMA B and the LLAMA F test results ($r = 0.447$; $p = 0.168$) as evident in Fig. 2 above. This means when females scored high on one test, it was likely they also scored high on the other test.

4 Discussion

The purpose of this study was to investigate whether there are differences between males and females in both grammar and vocabulary acquisition. The results obtained in this study cannot support the claim that there are differences between the genders. H1, which assumed that females perform better than males at learning unknown words and therefore score higher on the LLAMA B test, could not be supported. Likewise, H2 2, which expected that males are better at learning grammatical rules than females and score higher in the LLAMA F test, could not be supported.

H1 assumed that females would outperform males with regard to vocabulary learning due to many other studies reporting that women have superior declarative memory ability (Ullman et al., 2008) and should therefore be better at learning unknown vocabulary. For example, Kimura (1999) reported that females perform better in episodic memory tasks, such as remembering a given set of stimuli (e.g. words, digits, paragraph content, nameable items and real objects). Additionally, Huttenlocher et al. (1991) and Horgan (1975) tested children between 2 and 4 years of age and found that females showed a larger vocabulary repertoire and produced more complex utterances. Furthermore, Kimura (1999), Lewin et al. (2001) and Saucier et al. (2003) all ascribe females' superior memory for non-verbal stimuli, such as faces and objects, and their reliance on landmarks during spatial navigation tasks to their ability to internally verbalise these stimuli better than males. Even though the mean results of the LLAMA B test show that females score slightly better than males, which is in accordance with the literature results, the further statistical analysis yielded no significant difference between the genders ($p = 0.870$). The interpretation of the results is difficult because of the small number of participants ($n = 18$). Nonetheless, there is the possibility that what is suggested by the mean values could become a significant difference if more participants were investigated.

H2 expected that males would outperform females in grammar learning based on findings which demonstrate that men have superior procedural memory ability and should therefore be better at grammatical rule learning (Hartshorne & Ullman, 2006). Albeit this study did not find any statistically significant results, the results of the t-test for the LLAMA F were closer to the statistically significant threshold compared to the LLAMA B ($p = 0.146$), and would almost certainly pass it if more participants were added to the study. Hence, the result could be seen as a cautious trend toward a difference between males and females in the LLAMA F test. Notably, a trend could also be observed in the descriptive statistics considering the mean values of both genders in this test as females had on average higher scores than males. As Ullman et al., 2008; Ullman et al., 2002; Ullman (2004, 2005) and Hartshorne and Ullman (2006) extensively describe, males are thought to have superior procedural memory abilities and are consequently deemed to be better at learning grammatical rules. Due to these numerous pieces of evidence in the literature, it can be speculated that the outcome of testing the hypothesis of this study was indeed influenced in favour of the male participants concerning the LLAMA F test, which is supposed to test for grammatical rule learning ability.

Interestingly, a study by Wucherer and Reiterer (2016) found a superiority of females in grammar learning. The results of this study might therefore indicate that females actually are better at learning grammatical rules than males, and that the inferences from superior male spatial and navigational skills, which Ullman (2004) draws, are not applicable to grammar learning. In fact, several studies present controversial results concerning this topic. For instance, McGivern et al. (1997) found that both girls and women were better at remembering complex abstract patterns than their male counterparts. These patterns can also be seen as a sequence of abstract stimuli and from the literature Ullman et al. (2008) presented in his review, male participants should have performed better due to their superior procedural memory skills.

This description of females having a superior memory for non-verbal stimuli like faces and objects due to their ability to internally verbalise these stimuli has been mentioned in the literature (for example Ullman et al., 2008). Interestingly, the most successful females in both the LLAMA B and LLAMA F test reported in the post-testing interviews that they were able to establish emotional connections between the objects and the names they were supposed to remember. For example, 2 females reported that they were easily able to remember one object in the LLAMA B test, which was called ‘CIMI’, because it reminded them of a cat and the name ‘Jimmy’. They further reported that they remembered all objects in a similar way and repeated this connection in their minds (‘Jimmy the cat’) several times. Thus, they used internal verbalisation and scored between 95 and 100 points. This ties in with the aforementioned literature and is interesting in several regards. For instance, Lewin et al. (2001) and Saucier et al. (2003) reported that females’ superior declarative memory skills are tied to their ability to internally verbalise these stimuli and that this system may be particularly important for learning idiosyncratic information, specifically arbitrary relations.

In a study by Kansaku et al. (2000) only females showed bilateral activation in brain areas when listening to a story in a functional MRI scanner. Interestingly, Baron-Cohen et al. (2005) states that connectivity in the brain is linked to the ability to empathise. Therefore, the fact that females show interhemispheric activity when confronted with language related tasks could imply an emotional involvement in declarative memory retrieval.

Ullmann et al., (2008) also claim that declarative memory ability links to spatial navigation tasks, and thus explains why women orientate themselves by landmarks and men rather by spatial cues, such as depth and sequences.

Prominently, the empathising-systemising (E-S) theory (Baron-Cohen et al., 2005) ties in with the declarative/procedural memory system. Interestingly, when comparing the two dimensions, males are generally thought to be more systemising (trying to find patterns and rules) and females more empathising (trying to understand others’ emotional state). Therefore, the fact that the most successful female participants reported that they remembered stimuli by forming explicitly emotional connections to them (‘CIMI looks like my cat Jimmy’; ‘This one looks angry and is called MEN, so he is an angry man’) offers strong evidence that they used the empathising memory system described by Baron-Cohen et al. (2005).

Notably, none of the male participants of this study reported using a strategy similar to internal verbalisation or establishing an emotional link for remembering the objects in the LLAMA B test. Most male participants, but also some females, reported to have tried to remember the sequence of objects or the location on the screen, which corresponds to the systemising memory dimension according to E-S theory (Baron-Cohen et al., 2005). Apparently though, looking at the test scores none of these strategies were as successful as the emotional connection. Particularly interesting with regard to males trying to remember objects by sequence or location is the fact that according to Williams et al. (1990), Bever (1992), Sandstrom et al. (1998) and Saucier et al. (2002), males are thought to be better at tasks that require spatial skills and navigating directions, because they use the left hippocampus for recognising geometrical cues, spatial depth and visuospatial processing. These observations are certainly interesting; however, further studies with more participants are needed to shed more light on them.

5 Conclusion

As discussed in the previous section, the current study yielded no significant results; however, potential reasons for this outcome can help researchers design a better suitable study in the future. This study had its limitations and several explanations for why the results in this paper contradict the literature come to mind. First and foremost, the number and selection of participants has to be seen critically. The small sample size has to be considered especially where males are concerned ($n = 7$), because the explanatory power and significance of the results have to be treated with special caution. However, even with a larger sample size this study would have its limitations.

For instance, even if the results show a slight trend toward females outperforming males, in a larger and more balanced sample where education is controlled for (only 3 BA and MA students in males, but 9 BA and MA students in females), the results might become significant in the opposite direction. According to the design, the LLAMA tests are education, language and age free (LLAMA Language Aptitude tests); however, participants with a tertiary level of education might have advantages with regard to learning strategies.

Another explanation for the insignificant outcomes could be the design of the LLAMA tests themselves. As the authors of the test explicitly state on their website, the LLAMA tests were developed as an alternative to other language aptitude tests restricted to a certain language, and are still in their testing phase. It is possible that due to the fact that the LLAMA F also entails remembering different pictures and shapes, and connecting them to invented words, similar mechanisms as in the LLAMA B might be at work. Consequently, both tests might actually test the same feature, namely vocabulary learning. A result that might support this claim is a positive correlation between the test scores of the LLAMA B and the LLAMA F test in females. Women, who scored high on one test, also scored high on the other, which might indicate a similarity in their task set-up and that the LLAMA B and LLAMA F actually test similar abilities. On the other hand, it might be the case that females who scored high on one test also scored high on the other because they were simply good at learning both, vocabulary and grammar rules (see Wucherer & Reiterer, 2016).

To summarise, several studies show that women have greater interhemispheric activity than men in language related tasks, pointing towards a stronger emotional involvement in declarative memory retrieval, while the same process in men is performed by only one hemisphere connected to the procedural memory system, specialised for rules and sequences. This study investigated whether women would outperform men in vocabulary memory tasks as tested by the LLAMA B, and whether men would perform better in grammar learning as tested by the LLAMA F. These hypotheses were formulated due to the assumption that men rely more on the procedural memory system for memorising lexical information, while females rely on declarative memory for the same task. Even though the results were not statistically significant, indications for a trend for a difference between men and women was found in the LLAMA F test. Due to the small sample size of the study, it would be inappropriate to draw definite conclusions, but it could give fresh impulses to the research in this field nonetheless. For instance, future researchers could extend the post-test interviews about the individual memory strategies, as it revealed some of the most interesting points of this study, but was unfortunately not recorded systematically. This qualitative information in combination with the quantitative test results could enable new insights and a more grounded assessment of the nature of gender differences in vocabulary and grammar learning.

References

Alexander, G. M., Packard, M. G., & Peterson, B. S. (2002). Sex and spatial position effects on object location memory following intentional learning of object identities. *Neuropsychologia, 40*(8), 1516–1522. https://doi.org/10.1016/S0028-3932(01)00215-9

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: Implications for explaining autism. *Science, 310*(5749), 819–823. https://doi.org/10.1126/science.1115455

Baxter, L. C., Saykin, A. J., Flashman, L. A., Johnson, S. C., Guerin, S. J., & Babcock, D. R. (2003). Sex differences in semantic language processing: A functional MRI study. *Brain and Language, 84*(2), 264–272. https://doi.org/10.1016/S0093-934X(02)00549-7

Bever, T. G. (1992). The logical and extrinsic sources of modularity. In M. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition* (pp. 179–212). London: Lawrence Erlbaum.

Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews Neuroscience, 7*(6), 477–484. https://doi.org/10.1038/nrn1909

Fuchs, R. A., Evans, A., Mehta, R., Case, J. M., & See, R. E. (2005). Influence of sex and estrous cyclicity on conditioned cue-induced reinstatement of cocaine-seeking behavior. *Psychopharmacology, 179*(3), 662–672. https://doi.org/10.1007/s00213-004-2080-7

Halpern, D. F., & Tan, U. (2001). Stereotypes and steroids: Using a psychobiosocial model to understand cognitive sex differences. *Brain Cognition, 45*(3), 392–414. https://doi.org/10.1006/brcg.2001.1287

Hartshorne, J. K., & Ullman, M. T. (2006). Why girls say 'holded' more than boys. *Developmental Science, 9*(1), 21–32. https://doi.org/10.1111/j.1467-7687.2005.00459.x

Hausmann, M. (2005). Hemispheric asymmetry in spatial attention across the menstrual cycle. *Neuropsychologia, 43*(11), 1559–1567. https://doi.org/10.1016/j.neuropsychologia.2005.01.017

Horgan, D. M. (1975). Language development. In *Neuroscience*. Ann Arbor: University of Michigan.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Kansaku, K., Yamaura, A., & Kitazawa, S. (2000). Sex differences in lateralization revealed in the posterior language areas. *Cerebral Cortex, 10*(9), 866–872. https://doi.org/10.1093/cercor/10.9.866

Kaufman, M. J., Levin, J. M., Maas, L. C., Kukes, T. J., Villafuerte, R. A., Dostal, K., et al. (2001). Cocaine-induced cerebral vasoconstriction differs as a function of sex and menstrual cycle phase. *Biological Psychiatry, 49*(9), 774–781. https://doi.org/10.1016/S0006-3223(00)01091-X

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT Press.

Lewin, C., Wolgers, G., & Herlitz, A. (2001). Sex differences favoring women in verbal but not in visuospatial episodic memory. *Neuropsychology, 15*(2), 165–173. https://doi.org/10.1037//0894-4105.13.4.590

Llama Language Aptitude tests [Computer Software]. (2015). Retrieved from http://www.lognostics.co.uk/tools/llama/

McGivern, R. F., Huston, J. P., Byrd, D., King, T., Siegle, G. J., & Reilly, J. (1997). Sex differences in visual recognition memory: Support for a sex-related difference in attention in adults and children. *Brain and Cognition, 34*(3), 323–336. https://doi.org/10.1006/brcg.1997.0872

McGlone, J. (1977). Sex differences in the cerebral organization of verbal functions inpatients with unilateral brain lesions. *Brain, 100*(4), 775–793.

Phillips, S. M., & Sherwin, B. B. (1992). Effects of estrogen on memory function in surgically menopausal women. *Psychoneuroendocrinology, 17*(5), 485–495.

Sabbatini, R.M.E. (1997). *Are there differences between the brains of males and females?*. Retrieved from http://www.cerebromente.org.br/n11/mente/eisntein/cerebro-homens.html

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research, 6*(4), 351–360.

Saucier, D., Bowman, M., & Elias, L. (2003). Sex differences in the effect of articulatory or spatial dual-task interference during navigation. *Brain and Cognition, 53*(2), 346–350. https://doi.org/10.1016/S0278-2626(03)00140-4

Saucier, D., Green, S., Leason, J., MacFadden, A., Bell, S., & Elias, L. J. (2002). Are sex differences caused by sexually dimorphic strategies or by differences in the ability to use the strategies? *Behavioral Neuroscience, 116*(3), 403–410. https://doi.org/10.1037//0735-7044.116.3.403

Schlaepfer, T. E., Harris, G. J., Tien, A. Y., Peng, L., Lee, S., & Pearlson, G. D. (1995). Structural differences in the cerebral cortex of healthy female and male subjects: A magnetic resonance imaging study. *Psychiatry Research, 61*(3), 129–135.

Shaywitz, B., Shaywitz, S. E., Pugh, K. R., Constable, R. T., Skudlarski, P., Fulbright, R. K., et al. (1995). Sex differences in the functional organization of the brain for language. *Nature, 373*(6515), 607–609. https://doi.org/10.1038/373607a0

Sherwin, B. B. (1988). Estrogen and/or androgen replacement therapy and cognitive functioning in surgically menopausal women. *Psychoneuroendocrinology, 13*(4), 345–357.

Stansfield, C. W. (1989). Review of the Pimsleur language aptitude battery. In D. J. Keyser & R. C. Sweetland (Eds.), *Test critiques* (Vol. III, pp. 438–445). Kansas City, MO: Test Corporation of America.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231–270.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language* *acquisition: Methods, theory and practice* (pp. 141–178). Washington, DC: Georgetown University Press.

Ullman, M. T., Estabrooke, I. V., Steinhauer, K., Brovetto, C., Pancheva, R., & Ozawa, K. (2002). Sex differences in the neurocognition of language. *Brain and Language, 83*(1), 141–143.

Ullman, M. T., Miranda, R. A., & Travers, M. L. (2008). Sex differences in the neurocognition of language. In J. B. Becker, K. J. Berkley, & N. Geary (Eds.), *Sex on the Brain: From genes to behavior* (pp. 291–309). New York/Oxford: Oxford University Press.

Williams, C. L., Barnett, A. M., & Meck, W. H. (1990). Organizational effects of early gonadal secretions on sexual differentiation in spatial memory. *Behavioral Neuroscience, 104*(1), 84–97.

Wucherer, B., & Reiterer, S. M. (2016). Language is a girlie thing, isn't it? A psycholinguistic exploration of the L2 gender gap. *International Journal of Bilingual Education and Bilingualism, 19*, 1–17. https://doi.org/10.1080/13670050.2016.1142499

Part IV
Language Aptitude and Socio-environmental Influences

Vocabulary Acquisition Strategies & Language Aptitude

Jakob Poschner

Abstract According to the lexical approach in language teaching, it is lexis which functions as the most important building block for communication. Therefore, it is of great importance to examine how vocabulary, also in the form of multiword lexical units, can be acquired most efficiently. Vocabulary learning strategies haven been a constant field of research since the 1970s, providing linguists and teachers with information about different cognitive vocabulary learning strategies. Additionally, language aptitude tests improved significantly since their development and were able to prove their validity. Studies showed that low aptitude learners drastically improved with the help of specific vocabulary acquisition strategies, whereas high aptitude learners did not benefit from such strategies. Taking this finding into consideration, this research investigates whether people with a lower language aptitude use particular cognitive strategies more frequently compared to persons with a higher aptitude. A questionnaire has been designed to measure people's preferences for various cognitive vocabulary acquisition strategies. The questionnaire focuses on three well known and two little-known, but innovative, strategies: learning vocabulary with the help of pictorial representations, grouping corresponding words together, antonyms and synonyms, using physical actions accompanying vocabulary learning and the keyword method. A total of 19 German native speakers participated in the study. The LLAMA B test has been chosen to measure vocabulary acquisition aptitude. The research shows that low aptitude learners do not use significantly more cognitive vocabulary acquisition strategies than high aptitude learners.

J. Poschner (✉)
Department of Linguistics & Department of English and American Studies, University of Vienna, Vienna, Austria
e-mail: a0748537@unet.univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_13

1 Introduction

According to the lexical approach in language teaching, it is lexis which functions as the most important building block for communication and not grammar, functions, notions, or a different area of teaching and planning (Richards & Rodgers 2001). Therefore, it is of great importance to examine how vocabulary, also in the form of multiword lexical units, can be acquired in the most efficient way, and in how far this relates to language aptitude.

In order to offer a context to the reader, this paper provides introductory information based on insights gained by Markus Hengstschläger (2012) with regard to talents and genes. Since language aptitude tests are concerned with the notion of talent, it seems important to provide the state of the art of this concept in Sect. 1.1.

Vocabulary learning strategies haven been a constant field of research since the 1970s, providing linguists and teachers with information about different cognitive vocabulary learning strategies. Additionally, language aptitude tests improved significantly since their development and were able to prove their validity in numerous studies (Granena & Long 2013; Ranta 2002; Doughty et al. 2010). Since research was able to demonstrate that people display different levels of language aptitudes, the necessity to take this into account in vocabulary learning and teaching becomes evident. However, scarcely any studies have established the connection between second language aptitude and vocabulary learning strategies. One of the few research papers to do so were conducted by Wieland (1990) and Macedonia, Müller, and Friederici (2010). Wieland and Macedonia, Müller and Friederici divided the participants of their studies into groups. The different groups studied vocabulary items either with or without a specific vocabulary acquisition strategy. In Macedonia, Müller and Friederici's case the participants used gestures accompanying vocabulary acquisition; in Wieland's case subjects learned with the help of the keyword method. Prior to the vocabulary learning phase all subjects were pretested with a language ability test in order to later determine a possible correlation between vocabulary recall performance and language aptitude scores. Both studies reached the same conclusion, namely that people who easily acquire new vocabulary do not improve with the help of a specific strategy; however, persons who have difficulties learning vocabulary improve drastically. A detailed explanation of the two studies follows in Sect. 1.4.

In order to further close the gap between language aptitude findings and vocabulary acquisition techniques, this research investigates whether individuals with a lower language aptitude use particular strategies more frequently compared to persons with a higher aptitude. For this purpose, this paper will focus on three well known and two little-known, but innovative, cognitive vocabulary acquisition strategies, namely learning vocabulary with the help of pictorial representations, grouping corresponding words together, antonyms and synonyms, using physical actions accompanying vocabulary learning and mnemonic techniques, such as the keyword method (Schmitt & McCarthy 1997).

This paper is expected to yield novel insights into the actual learning behavior of foreign language learners by investigating the following research question: Do people with a low aptitude for vocabulary learning incorporate cognitive vocabulary acquisition strategies more frequently than people who learn new words with ease? This question provides further grounds for analyses which depend on its findings. For instance, if low aptitude learners do not implement strategies which can improve their vocabulary learning drastically, the question of how to increase their motivation to use them arises. However, if they already make use of these strategies it would be necessary to investigate whether certain strategies prove to be especially useful for low language aptitude learners.

1.1 Genes and Talent

Markus Hengstschläger, professor of medical genetics and leader of the fundamental research program at the University of Vienna, sums up the state of the art in science with regard to the interplay between genes and talent. With respect to nature or nurture, Hengstschläger (2012) holds the opinion that it should rather be called nature and nurture, although there clearly are genetical attributes which predefine the potential for performance. For example, genes are responsible for the anatomy of the vocal organ, which then again is responsible for the particular sound of a voice. Furthermore, certain sports such as basketball and volleyball require the athletes to acquire a genetically influenced feature, namely a certain height (Hengstschläger 2012). On the other hand, talent is developed through practice and the influence of the environment. According to Hengstschläger (2012), it was Anders Ericsson in the 1990ies who empirically proved that all outstanding performances were the result of years of regular practice. Additionally, many authors concluded that all it takes to excel in a certain area is a plethora of practice, oftentimes speaking of 10.000 h (Hengstschläger 2012). They resume that a person who outperforms another simply spent more time practicing. However, Hengstschläger (2012) raises serious reservations against this. He describes experiments with monozygotic twins, some of them being raised in the same family, whilst others were raised separately, with the aim to investigate the influence of both genes and the environment on the evolution of certain features. The experiments proved that numerous traits are inherited; however, it could also be shown that other components are strongly influenced by the environment. Hence, the conclusion drawn by Hengstschläger (2012) is that genes predispose certain talents, but that they always need practice in order to deliver outstanding performances. Since genes can affect talent, meaning that talent is to some extent inherited, the following question arises: In how far can talent be measured with regard to language learning? In order to answer this question, the following section provides a brief summary of the state of the art in the field of language aptitude testing.

1.2 Language Aptitude Tests

This section will describe the concept of language aptitude tests and briefly present its ongoing debate amongst linguists. Language aptitude tests attempt to predict the success in language learning by means of testing specific abilities (Lightbown & Spada 2013). According to Granena (2014) they comprise "cognitive and perceptual abilities that predispose individuals to learn well or rapidly" (p. 483). In addition, she states that there is an agreement on the importance of aptitude regarding second language acquisition. However, Lightbown and Spada (2013) fail to see this agreement. They claim that the considerable number of research pointing to a correlation between language aptitude and foreign language learning were conducted at a time when language teaching was primarily concerned with grammar translation and audiolingual methods. Lightbown and Spada (2013) argue that many researchers, albeit not giving any names, believe that the current communicative approach in teaching makes the abilities tested by language aptitude tests irrelevant to the process of language acquisition. In contrast to this view, Ranta (2002) demonstrated that children who scored high in language aptitude tests were the most successful learners in an environment which did not explicitly focus on grammar. A further study which shows the validity of such tests was conducted by Doughty et al. (2010), who were able to demonstrate the predictive power of an aptitude test via successful prognosis of high-level second language attainment. Therefore, the use of a language aptitude test seems justified for the purpose of this study.

1.3 Vocabulary Learning Strategies

Schmitt and Mccarthy (1997) divide language learning strategies into three major areas: the metacognitive area, the social/affective area, and the cognitive area. This research focuses on five different cognitive strategies for vocabulary acquisition, which manipulate information in order to acquire or retain that information. Three of those cognitive strategies are rather well-known since they are often encountered in text books. These are: to study a word with the pictorial representation of its meaning, to connect the word to its synonyms and antonyms, and grouping words together to study them. Notably, numerous studies indicate the benefits of such vocabulary learning strategies. For instance, studying new words with their pictorial representation of their meaning has been shown to be more successful than learning them with their direct translation in both Russian (Kopstein & Roshal 1954) and Indonesian (Webber 1978). Moreover, sense relationships such as synonyms (joy – happiness) and antonyms (bright – dark) can help students consolidate vocabulary (Schmitt & McCarthy 1997). Further, Cofer, Bruce, and Reicher (1966) state that when words are grouped together into meaning categories, for example, animals in one group and names in another group, their memorization is greatly improved.

Moreover, this paper concentrates on strategies which presumably are less frequently encountered in foreign language learning, such as memory strategies and using a physical action when learning a new word. Specifically, the foci lie on both the Voice-Movement-Icon Approach and the keyword method, which will be described in the subsequent paragraphs.

In terms of physical actions accompanying vocabulary learning in order to facilitate acquisition and retention, the Voice-Movement-Icon approach (VMI) embodies a characteristic example. Macedonia, who is an expert in the field of applied linguistics, has been an advocate for the VMI approach for over a decade. A VMI "consists of a word that is read and spoken in L2 [second language] and synchronously paired with an action or a gesture. A VMI is first performed by the language trainer and then imitated by the learners" (Macedonia 2013, p. 103). The multimodal acquisition with many different senses is the main benefit of a VMI. Words are perceived visually (students read the word, they see the VMI), auditorily (students hear the word pronounced by the teacher's, colleague's and their own mouth), kinesthetically (the VMI is reproduced by the students), in a proprioceptive way (the students feel their own movements), articulatorily (students say the word out loud), semantically (the VMI imitates the semantic content of the word, and the teacher explains the meaning of the word), and emotionally (VMIs arouse emotions since they are not common, oftentimes surprising, and amusing; Peßensteiner 2014). Numerous studies support the benefits proclaimed by VMI advocators. Macedonia (2013), for example, proves that nouns which are spoken and simultaneously accompanied by a gesture are retrieved better than when solely audio-visually learned. Furthermore, Macedonia, Müller & Friederici (2011) demonstrate that enactment positively affects abstract words' retrieval, and facilitates sentence production and retention over time (Macedonia 2013). Further research papers supporting the idea of using gestures for vocabulary acquisition were conducted by Tellier (2008), and Kelly, McDevitt, and Esch (2009), all of whom find that gestures lead to an improved retrieval for foreign words and phrases.

In terms of memory strategies, the keyword method has attracted considerable attention in foreign language studies. Thompson (1987) explains that the technique works by "utilizing some well-known principles of psychology: a retrieval plan is developed during encoding, and mental imagery, both visual and verbal, is used" (p. 66). In order to study the new word, a keyword in the first language is generated which resembles the new foreign language word in appearance or sound. Subsequently, an interactive image is developed in the mind which contains the keyword and the meaning of the new word (Wyra, Lawson, & Hungi 2007). For instance, the German word *Raupe* could be learned with the English keyword *rope*, which is similar in sound. The mental image would show a caterpillar (which is the translation of *Raupe*) lying on a *rope*. Hulstijin (1997) concludes that, on one hand, the keyword method should have its place in the classroom alongside contextual methods since its effect has been sufficiently proven, and it can be helpful for words which have not been successfully acquired. On the other hand, Hulstijn (1997) also claims that the keyword method can only be used for words which refer to objects that can be perceived visually and that it is less effective for backward recall. However, serious doubts can be raised against this. Wyra et al. (2007) point out that both backward and forward recall performance improve with the keyword method as long as attention is given to the keyword procedure at the time of retrieval. In fact, this improvement was consistent across all word classes.

1.4 Language Aptitude and Vocabulary Acquisition Research

The research conducted by Macedonia et al. (2010) investigated whether students learn new vocabulary better when paired with gestures that resemble the words' meanings than with arbitrary gestures. The subjects consisted of thirty-three native German speakers, 17 of which were female and 16 male, which were trained in two separate groups. With the aim to evaluate the participants' short-term memory and their ability to study novel words a pretest, based on a German nonword repetition task, was administered. Over the course of 4 days, 92 vocabulary items of an artificial corpus, Vimmi, were practiced in 4 sessions per day. Each session lasted 29 min and contained 23 words; thus, each block was repeated four times within 4 days. Performance was assessed daily from the second day on by translation of randomized lists from German into Vimmi and Vimmi into German. Subsequently, participants were divided into a high performers' group and a low performers' group, according to their vocabulary retrieval and production results. It was shown that high performers gained only minimal benefits from the accompanying corresponding gestures whereas low performers achieved considerable improvements. No correlation between the pretests' results of students' ability to learn new words and the vocabulary tests' results could be observed.

Wieland (1990) thoroughly examined variations of the keyword method in order to find an option which facilitates backward recall. Participants were 80 English speaking university students, 19 of which were male and 61 female, from 18 to 42 years of age. Similarly to the study conducted by Macedonia et al. (2010) the target vocabulary consisted of pseudo words. Each word was six letters long, presented on-screen for 10 s with its phonetical realization and was given with its English translation. One keyword group was provided with a concrete noun keyword which resembled the first three letters of the pseudo target word while another keyword group was asked to invent additional keywords for the second halves of the pseudo words. Control groups were not given any keywords, nor have they been informed about the keyword method. A language aptitude test was used in order to measure the participant's ability to identify familiar words as keywords for unfamiliar words. For this purpose, the third part of the MLAT, Modern Language Aptitude Test (Carroll & Sapon 1959), was chosen. It is concerned with Spelling Clues, 50 words are phonetically spelled and the subjects' aim is to decipher each word and find a synonym out of five possible alternatives. According to Dinklage (1971) it has been shown that it is a valid test for language aptitude, particularly for the ability to associate between sound and symbol (Gliksman, Gardner, & Smythe 1979). The research concluded that the results of the MLAT correlated positively with the vocabulary outcomes and that participants with a low aptitude score improved in backward vocabulary recall with the help of the keyword method.

1.5 Research Question

Taking the mentioned findings in the fields of language aptitude and vocabulary acquisition into consideration and based on studies which have shown that particular cognitive vocabulary acquisition strategies drastically improve the recall performance of people who have a low aptitude for vocabulary learning, this research attempts to answer the question whether these findings are represented in the way people with low vocabulary aptitude study vocabulary. The research question is: Do people with a low aptitude for vocabulary learning incorporate cognitive vocabulary acquisition strategies more frequently than people who learn new words with ease?

2 Methods

2.1 Participants

A total of 19 German native speakers participated in the study which covered the age range from 18 to 30 years. 32% were male (n = 6) and 68% were female (n = 13). Except for two individuals (10%), everybody had at least a high-school diploma, and three participants held a university degree (15%). The majority of participants (n = 14) were studying a language at the time of the survey or have been studying a language at university. Likewise, the majority of subjects (n = 10) were fluent in three languages.

2.2 Materials and Procedure

Participants were required to complete a language aptitude test. The test used for this research was the LLAMA B, which tests vocabulary learning aptitude. It is part of the LLAMA test battery (Meara 2005), which is the latest version of Swansea's University's language ability tests (Meara, Milton, & Lorenzo-Dus 2003). The test has an exploratory nature and is automatically scored. It is based on visual stimuli and verbal items developed out of a British-Colombian indigenous language and a

Central American language. It is, therefore, language independent. The study phase lasts 2 min; afterwards, the participants are asked to respond to a series of items without any time pressure (Granena 2014). The main reason why this test was chosen for this research was its user-friendliness and the circumstance that it was publically available. At the beginning, however, it appeared as if the test had a possible shortcoming, namely that it had not yet been extensively standardized (Granena & Long 2013). Since standardization requires the administration of a test to a large sample of various backgrounds, the LLAMA tests "should not be used in high-stakes situations" (Granena & Long 2013, p. 113). Granena and Long (2013), however, performed an exploratory validation study with the aim of assessing the reliability of the LLAMA with regard to its internal consistency and stability over time. Their conclusion was that the tests' internal consistency and stability over time is acceptable (Granena & Long 2013). Hence, although the test is not yet standardized, its use in numerous second language acquisition studies, and particularly in this research, is justified.

Subjects were also required to fill out an online questionnaire,[1] which consisted of three sections. In addition to basic questions concerning age, gender, education, and language fluency, detailed questions about participants' choices of vocabulary acquisition strategies were asked. A multi-item scale was used and items were grouped into five different concepts, namely the use of synonyms and antonyms, grouping corresponding words together, studying a word with the pictorial representation of its meaning, use of mnemonic strategies and the possibility to answer that no particular strategy is used for vocabulary learning. The questions were of behavioral and attitudinal nature, such as:

- Ich bevorzuge Vokabel mit Hilfe von Bildern zu lernen. [I prefer to learn vocabulary with the help of pictures].
- Typisch für mich ist, dass ich Vokabel mit Synonymen (bedeutungsähnlichen Wörtern) und/oder Gegensätzen lerne. [It is typical for me to study vocabulary with synonyms (words nearly equivalent in meaning) and/or opposites].
- Ich verwende so gut wie nie eine besondere Methode/Technik beim Vokabel Lernen, ich lerne einfach mit der deutschen Übersetzung. [I hardly ever use a specific method/technique for vocabulary acquisition, I simply learn with the direct German translation].

The questions could be answered with: I completely agree, I predominantly agree, I rather agree, I rather do not agree, I predominantly do not agree, and I completely disagree.

Furthermore, it was asked whether participants are familiar with the Voice-Movement-Icon approach or a similar technique, how they would classify themselves in terms of learning styles, and whether they do believe in the notion of learning styles. The reliability analysis showed that no Cronbach's alpha score was below 0.84.

[1] Available on request

3 Results

3.1 LLAMA B Scores

Descriptive statistics for the LLAMA B test scores displayed the following: The average aptitude score was 69.84, the highest scoring participant obtained 95 points out of 100 possible points, and the lowest scoring participant obtained 30. The distribution of scores was normal according to the one-sample Kolmogorov – Smirnov test (p = .200).

Figure 1 displays the frequency of scores obtained by the participants.

As Fig. 1 clearly shows, most participants scored in between 65 and 90 points, thus explaining the high average score of 69.84. Five out of nineteen participants scored below 65 points, whereas 14 participants scored 65 points or higher. According to the developers of the LLAMA B test, scores in between 25–45 are average scores; therefore, most people are expected to score within this range. 50–70 is a good score and 75–100 is an extraordinary score (Meara 2005). Hence, the average score of 69.84 is outstandingly high; explanations for this high score will be sought in a following section.

3.2 Vocabulary Learning Strategies

In order to answer the research question whether there is a correlation between a low LLAMA B score and the selection of cognitive vocabulary acquisition strategies, a t-test for independent samples was used. Specifically, the LLAMA B scores were tested with five vocabulary learning strategies for a statistically significant

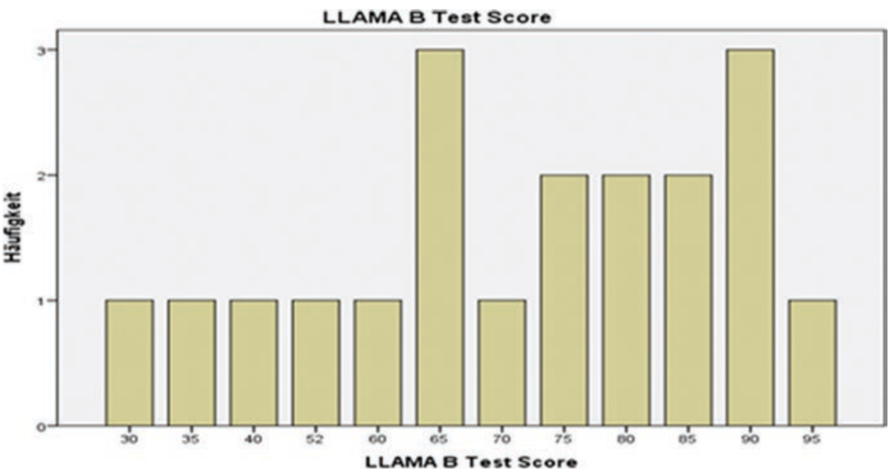


Fig. 1 Frequency of LLAMA B scores

difference. The vocabulary strategies were classified as mnemonic (representing mnemonics strategies in general), pictures (learning with pictorial representations), synonyms (the use of synonyms and antonyms), grouping (grouping words together into meaningful groups), and no technique (no specific strategy or technique is used). The LLAMA B scores were then divided into two groups, namely an average aptitude group (LLAMA B scores between 30 and 50) and a high aptitude group (LLAMA B scores between 55 and 95). Statistical differences were found for the groups synonyms ($t(17) = 2.12$, $p = .049$) and mnemonic ($t(17) = 2.19$, $p = .043$). The mean for the average aptitude group with regard to synonyms was 3.83 whereas it was 2.70 for the high aptitude group. A mean of 3.83 represents answers such as *I rather do not learn new vocabulary with the help of synonyms and antonyms* in the questionnaire. A mean of 2.70 represents answers such as *I rather do learn new vocabulary with the help of synonyms and antonyms*. A significant difference in means between the average group (mean = .433) and the high aptitude group (mean = .258) with regard to mnemonic strategies was found. Notably, there were no significant results obtained for both the high and average aptitude groups in the categories pictures, grouping, and no technique. The means in these groups ranged from 2,7 to 4,5 with the average being around 3; they reflect the rather equal distribution of people who use cognitive vocabulary acquisition techniques and people who do not. The Levene's test showed that the equality of variance assumption was met for all categories.

The significant results for the synonyms and mnemonic technique group have to be interpreted with caution since the average group consists of only three participants. Nonetheless, there seems to be a trend that people with a higher aptitude for vocabulary learning rather use mnemonic techniques and the help of synonyms and antonyms than people with a lower aptitude. In general, it can be said that the majority of students uses at least one cognitive vocabulary acquisition strategy.

Furthermore, the LLAMA B test scores were split into a very high aptitude group (> 75) and a high aptitude group (< 75). The idea behind this was to acquire an equally distributed number of subjects in each group, which was accomplished since the higher aptitude group consisted of 10 participants and the high aptitude group of 9 participants. However, no statistically significant difference could be found.

An ANOVA analysis was performed grouping the LLAMA B Scores into 3 concepts. An extraordinary high aptitude group with a LLAMA B test score of 75 and higher ($n = 10$), a high aptitude group with a score of 50 and higher ($n = 6$), and an average aptitude group with a score between 30 and 45 ($n = 3$). No statistically significant difference was noted among the groups.

3.3 Gender and LLAMA B Scores

A statistically significant difference was found between gender and LLAMA B scores. Figure 2 below shows the difference in score distribution between male and female participants.

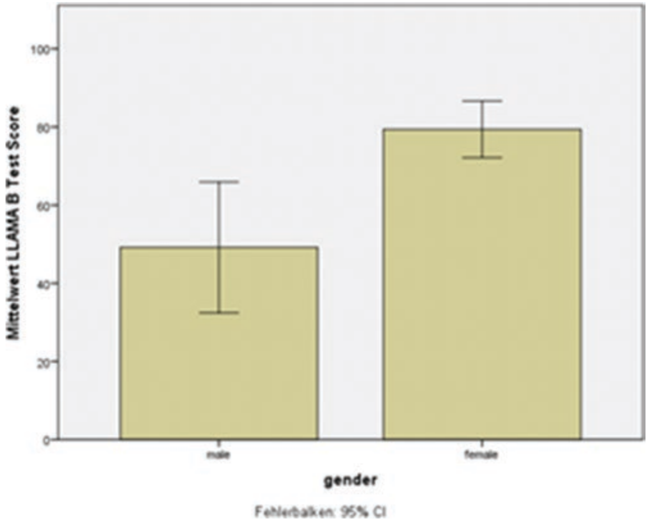


Fig. 2 T-test bar chart gender

As Fig. 2 demonstrates, female participants scored significantly higher in the LLAMA B test than their male counterparts. With respect to descriptive statistics it was found that 6 male participants reached an average score of 49.17 in the LLAMA B test; in contrast, female participants scored 79.38 points.

The t-test for independent samples clearly shows a strong, statistical difference between males and females ($p < .001$). The equality of variance assumption was met ($p = .134$) as the Levene's test demonstrated.

4 Discussion

The main aim of this study was to find an answer to the research question whether low aptitude vocabulary learners use cognitive vocabulary acquisition strategies more frequently. It was assumed that people who do not easily acquire new vocabulary, and therefore score low in a language aptitude test focusing on vocabulary such as the LLAMA B test, use vocabulary learning strategies in order to improve their vocabulary acquisition and retention. This speculation was based on research by Macedonia et al. (2010) and Wieland (1990), who both showed that people who have difficulties learning new vocabulary improved drastically with the help of particular vocabulary learning methods, whereas high aptitude learners did not improve. The data collected in this research with 19 participants did not show that people who scored low in the LLAMA B test use cognitive vocabulary acquisition techniques significantly more frequently than high aptitude learners. In contrary, the only two statistically significant differences in means with regard to vocabulary acquisition strategies suggest that people with a higher aptitude for vocabulary acquisition tend to use mnemonic techniques, synonyms, and antonyms for vocabulary acquisition more frequently. Mnemonic techniques, in this case, is an umbrella term for methods which require the learner to imagine a story or a picture associated with the new word in order to make associations with similar sounding words or written words, or to use rhymes for vocabulary acquisition and retention.

Since low aptitude vocabulary learners could improve drastically with the use of particular cognitive vocabulary learning strategies, the question of why they seemingly do not incorporate them more frequently than higher aptitude learners arises. A possible explanation could be that it takes time to practice and perfect a cognitive strategy such as the keyword method, or that organizing words together into meaning related groups proves to be too much effort. Likewise, there is the possibility that people are not aware of the different methods and the benefits which they can provide for them. However, it can be safely assumed that most participants in this study were aware of the majority of vocabulary strategies used for this research since 14 out of 19 subjects were studying a language at university, or learning it elsewhere, when the study was conducted. Thus, they are more likely to have knowledge of such strategies compared to people who are not interested in learning new languages.

When interpreting the LLAMA B scores, two insights were especially striking. Firstly, participants scored outstandingly high with a mean of 69.84. Secondly, women clearly outperformed men by scoring a mean of 79.38, whereas their male counterparts scored a mean of 49.17. These two findings need to be treated with caution though since the sample size for this research was rather small. Therefore, it can be argued that the results would differ with a larger sample size since such results are not typical when compared to LLAMA test results from other studies.

A comparison of the present study with the studies conducted by Macedonia et al. (2010) and Wieland (1990), which served as a starting point for the research question whether low aptitude vocabulary learners use more cognitive vocabulary acquisition strategies than their high aptitude counterparts, reveals both similarities and differences in: (1) the number of participants and gender distribution, (2) the subjects' language background and (3) the language aptitude test.

The study of Macedonia et al. (2010), Wieland (1990) and the one at hand vary considerably in the number of participants and in their gender distribution. Nineteen subjects, 6 male and 13 female, with a range from 18 to 30 years of age participated in the present study; Wieland's research (1990) counts 80 participants, 19 male and 61 female, ranging from 18 to 42 years of age, and the study by Macedonia et al. (2010) consisted of 33 participants, 16 male and 17 female, with a mean age of 23,17, no information about the age range could be obtained. The gender distribution shows that significantly more females participated in the studies with the exception of Macedonia, Müller and Friederici's research (2010), which has an equal distribution of male and female participants. Taking the present research's number of participants into account, it seems plausible to understand the result as an indication of a trend rather than a representatively proven state of affairs.

With regard to the participants' language background certain differences can be recognized among the presented studies. In Wieland's research 72 out of 80 undergraduate students were English native speakers; the rest was proficient in English as well. In Macedonia, Müller and Friederici's study (2010) all subjects were German native speakers. Both studies do not give information about the number of languages spoken by the participants nor about their history with foreign language learning. Equally to Macedonia, Müller and Friederici's research (2010), the participants in this study were all German native speakers. The circumstance that the majority of subjects in this research, 14 out of 19, were enrolled in a foreign language program at university or were learning a new language at the time of the survey is highly beneficial for the validity of the outcome. It seems reasonable to assume that these participants were familiar with the different cognitive vocabulary acquisition strategies, even more so because many subjects were English students at the University of Vienna, who are expected to keep a vocabulary log and are taught different strategies for that purpose.

Differences among the language aptitude tests can be observed in their selection while their purpose was the same. In Macedonia, Müller and Friederici's study (2010) the aptitude test was based on a German nonword repetition task (Gathercole 2006) with the idea to measure the phonological short term memory on the assumption of its importance for the learning procedure of novel words. Wieland (1990) used Part 3 of the Modern Language Aptitude Test (Carroll & Sapon 1959). The testing procedure contains the disguised spelling of a word; this word needs to be understood in order to choose the correct synonym out of 5 alternatives. According to the developers it is necessary to speak English for this task. The test was used to indirectly measure the participant's ability to successfully find keywords for a new word in their native language. In both studies, the aim of the aptitude tests was to clarify whether there is a correlation between the participants' aptitude test scores and their recall performance of the vocabulary learned with a specific cognitive vocabulary acquisition strategy. The aptitude test successfully correlated with recall performance in Wieland's research (1990) but failed to do so in Macedonia, Müller and Friederici's case (2010). Nonetheless, both studies showed the positive effect of specific vocabulary acquisition techniques for people who struggle with vocabulary retention.

The language aptitude test used for this study was the LLAMA B test; it was the most recent version of Swansea's University's language ability tests (Meara et al. 2003). Two features of the test are quite different from the tests used in the previously mentioned studies. First, it does not rely on phonetical input, participants receive visual stimuli and verbal items. Second, the language of the test attempts to be language independent; in contrast, Wieland's (1990) and Macedonia, Müller and Friederici's aptitude tests' languages were clearly based on the participants' native language. Interestingly, both research papers used a pseudo language for their vocabulary acquisition material but failed to do so in their aptitude testing. The reasons to test for an artificial corpus of language are to avoid associations with the target language or other languages known by the participants and to rule out other factors in natural languages that might influence participants' vocabulary learning. On a single word level these factors comprise "peculiarities in their phonotactics, word length, and frequency of occurrence" (Macedonia & Knösche 2011, p. 198).

Therefore, it seemed important for this study to choose a language aptitude test which is based on an artificial language. The test's purpose was to directly measure the participants' ability to memorize new vocabulary in order to divide them into groups of low performers and high performers.

5 Conclusion

Some results of this study should be interpreted with caution. A slight tendency of high aptitude vocabulary learners rather using synonyms, antonyms and mnemonic strategies was found. However, the statistical significance was close to being nonsignificant. Furthermore, the average group consisted of only three participants since, unexpectedly, the participants scored extraordinarily high in the LLAMA B test.

With regard to the main research question it can be safely stated that low aptitude vocabulary learners do not use cognitive vocabulary strategies in a different way than higher aptitude vocabulary learners. Numerous questions arise, such as, whether language learners are aware of such strategies' benefits, what reasons there are not to apply them more often, and in how far learners can be encouraged to apply them.

As a final point, a pedagogical implication will be given. If low aptitude students are not aware of the different vocabulary learning strategies and their benefits for them, then these methods should be taught in school. Not in order to obligate students to make use of a particular strategy but in order to make them familiar with various methods to personally experience possible benefits. Finally, after introducing such techniques, students need to be allowed to make their own decisions whether they want to apply such strategies or not.

References

Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test (Form A)*. New York: Psychological Corp.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. (1966). Clustering in free recall as a function of certain methodological variations. *Journal of Experimental Psychology, 71*, 858–866.

Dinklage, K. T. (1971). Inability to learn a foreign language. *Emotional Problems of the Student, 99*, 185–206.

Doughty, C. J., Campbell, S. G., Mislevy, M. A., Bunting, M. F., Bowles, A. R., & Koeth, J. T. (2010). Predicting near-native ability: The factor structure and reliability of Hi-LAB. In M. P. Prior, Y. Watanabe, & S. Lee (Eds.), *Selected Proceedings of the 2008 Second language research forum: Exploring SLA perspectives, positions, and practices* (pp. 10–31). Sommerville, MA: Cascadilla Proceedings Project.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied PsychoLinguistics, 27*, 513–543. https://doi.org/10.1017/S0142716406060383

Gliksman, L., Gardner, R. C., & Smythe, P. C. (1979). An investigation of the test-retest reliability of the subscales of the Modern Language Aptitude Test (MLAT). *Canadian Modern Language Review, 5*, 204–210.

Granena, G. (2014). Language aptitude and long-term achievement in early childhood L2 learners. *Applied Linguistics, 35*(4), 483–503. https://doi.org/10.1093/applin/amu013
Granena, G., & Long, M. (2013). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Philadelphia: John Benjamins Publishing Company.
Hengstschläger, M. (2012). *Die Durchschnittsfalle: Gene – Talente – Chancen*. Salzburg, Austria: Ecowin Verlag.
Hulstijn, J. H. (1997). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition, 19*, 131–143.
Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes, 24*(2), 313–334. https://doi.org/10.1080/01690960802365567
Kopstein, F. F., & Roshal, S. M. (1954). Learning foreign vocabulary from pictures vs. words. *American Psychologist, 9*, 407–408.
Lightbown, P. M., & Spada, N. (2013). *How languages are learned*. Oxford: Oxford University Press.
Macedonia, M. (2013). Learning a second language naturally the voice movement icon approach. *Journal of Educational and Developmental Psychology, 3*(2), 102–116. https://doi.org/10.5539/jedp.v3n2p102
Macedonia, M., & Knösche, R. T. (2011). Body in mind: How gestures empower foreign language learning. *Mind, Brain, and Education, 5*(4), 196–211. https://doi.org/10.1111/j.1751-228X.2011.01129.x
Macedonia, M., Müller, K., & Friederici, A. D. (2010). Neural correlates of high performance in foreign language vocabulary learning. *Mind, Brain, and Education, 4*(3), 125–134. https://doi.org/10.1111/j.1751-228X.2010.01091.x
Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping, 32*(6), 982–998. https://doi.org/10.1002/hbm.21084
Meara, P. (2005). *LLAMA language aptitude tests*. Swansea: Lognostics.
Meara, P., Milton, J., & Lorenzo-Dus, N. (2003). *Swansea Language Aptitude Tests (LAT) v.2.0*. Swansea: Lognostics.
Peßensteiner, S. (2014). Wie bleiben die Wörter im Kopf? Gehirngerechter Wortschatzerwerb im Französischunterricht. Diploma thesis, University of Graz.
Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 159–180). Amsterdam/Philadelphia: John Benjamins.
Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge, UK: Cambridge University Press.
Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy*. Cambridge, UK: Cambridge University Press.
Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture, 8*, 219–235. https://doi.org/10.1075/gest.8.2.06tel
Thompson, I. (1987). Memory in language learning. In A. Wenden & J. Rubins (Eds.), *Learner strategies in language learning*. New York: Prentice Hall.
Webber, N. E. (1978). Pictures and words as stimuli in learning foreign language responses. *The Journal of Psychology, 98*, 57–63.
Wieland, L. R. (1990). The effects of training and the keyword method on the recall of an unfamiliar vocabulary. PhD thesis, University of Ottawa.
Wyra, M., Lawson, M. J., & Hungi, N. (2007). The mnemonic keyword method: The effects of bidirectional retrieval training and of ability to image on foreign language vocabulary recall. *Learning and Instruction, 17*(3), 360–337. https://doi.org/10.1016/j.learninstruc.2007.02.008

Comparing the Language Aptitudes and Language Attitudes of Mono- and Bilingual Burgenland Croats

Katharina Krumpeck

Abstract Research on bi- and multilingualism in connection to language aptitude has primarily shown that they are mostly positively related. Yet, learning a foreign language is not only associated with a person's language aptitude, but language attitudes can also play a vital role. Having a high language aptitude does not necessarily result in acquiring a language more easily than someone who is highly motivated to learn a specific language and whose language attitudes regarding that tongue are positive. Thus, it seems reasonable to investigate both language aptitudes and language attitudes, especially of a minority group such as Burgenland Croats, as it could be argued that their minority language, Burgenland Croatian, is facing language death. Given previous research on bi- and multilingualism, it was assumed that bilingual Burgenland Croats have a higher language aptitude than monolingual Burgenland Croats. Additionally, because of their assumed closer affiliation with Burgenland Croatian and its corresponding culture and tradition, it was deduced that they have a more positive attitude towards their minority language. These hypotheses were tested by means of the MLAT Part V, as well as the LLAMA_B (online version) and an attitudes questionnaire. Calculations were done by using IBM SPSS Statistics (Version 23). The research has shown that bilingual Burgenland Croats have a more positive attitude towards their minority language.

1 Introduction

The following section will provide a general overview on bilingualism, language aptitude, language attitudes, language shift and language death. Aside from offering definitions, studies from previous research on bi- and multilingualism, and language aptitude and attitudes will be evaluated.

To start with bilingualism, "[t]he most common definition […] is the ability to speak two languages" (Korani, 2012, p. 1747). This is in concordance with

K. Krumpeck (✉)
Department of Linguistics, University of Vienna, Vienna, Austria

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_14

Weinreich's (1966) rather broad definition of bilingualism, namely "the practice of alternately using two languages" (p. 74). Jessner (2008), like many other researchers, believes "that it is best to view bilingualism on a continuum" (p. 20). This entails that there are many variances as to what can be classified as being a bilingual. Planchon and Ellis (2014), for instance, understood bilinguals as

> [people] who consider themselves to 'speak' more than one language to the extent that they can use them to achieve communicative ends, including in limited circumstances, or whose responses indicate that they are capable of doing this, even if they do not consider themselves to be so. (p. 211)

However, what *achieving communicative ends* exactly means is rather vague and unclear, as it is very subjective and probably differs from individual to individual. A closer gradation within bilingualism is provided by Costa and Sebastián-Gallés (2014) who differentiate between simultaneous and successive bilinguals. *Simultaneous bilinguals*, in their opinion, "learn two languages from birth" (Costa & Sebastián-Gallés, 2014, p. 336), however, they expand the meaning of the term by declaring that it "is also used to refer to individuals who acquire their second language within the first 2–3 years of life" (Costa & Sebastián-Gallés, 2014, p. 337). *Successive bilinguals*, on the other hand, "learn a second language later in life under formal instruction, in an immersion environment as a result of immigration or in one of many other situations […]" (Costa & Sebastián-Gallés, 2014, p. 336). These definitions enable a fairly clear and helpful way of distinguishing between different levels of bilingualism and allow for a circumnavigation of the troublesome way to define bilingualism by means of *native-like proficiency*. Related to this important aspect of native-like proficiency of bilinguals is dominance. Costa and Sebastián-Gallés (2014) point out that studying bilinguals is a rather difficult task insofar as speakers' "proficiency levels" may differ highly from one another, namely that "whereas some show equal proficiency in both languages, others have a clear dominance in one of their languages" (p. 336).

A majority of research on bilingualism compares mono- and bilinguals in order to identify the potential advantages of bilingualism as opposed to monolingualism. However, Costa and Sebastián-Gallés (2014) clarify that "the neural networks involved in first-language processing seem to be fundamentally the same for monolinguals and bilinguals […]" (p. 343). Nevertheless, even though both groups seem to have the same starting point, they note that due to "fac[ing] higher processing demands that lead to an increase in brain activity", bilinguals may have an increased "cognitive reserve" in old age (Costa & Sebastián-Gallés, 2014, p. 343). However, Costa and Sebastián-Gallés (2014) also indicate that this advantage can only develop "if the linguistic input in any language (and particularly the first one) is frequent, varied and socially useful" (p. 343). If this is not the case, they argue that "one finds situations of switched language dominance in which the second language of a bilingual becomes the dominant language" (Costa & Sebastián-Gallés, 2014, p. 343). Such shifts from one language to another will be discussed in more detail in a subsequent paragraph. Yet, whereas Costa and Sebastián-Gallés (2014) propose that bilingualism can be advantageous, especially in old age, Jessner (2008) asserts that "[u]ntil the early 1990s, it seemed to be clear for most scholars that contact with more than one language would have to result in *problems* of either a cognitive or a linguistic nature" [my emphasis] (p. 17). Planchon and Ellis (2014), on the other hand, refer to researchers such as Liedtke and Nelson (1968), Peal and Lambert (1962) and Genesee, Tucker and Lambert (1975), who all noted advantages of bilingualism on children already at the beginning of the second half of the twentieth century. In their research on the effects of bilingualism, Planchon and Ellis (2014) focus in particular on *metalinguistic awareness* and assert that "metalinguistic understanding and cognitive flexibility" is particularly distinct with bilinguals and may facilitate the successful acquisition of additional languages (p. 203). Bialystok (2001), being critical of the increased "metalinguistic ability *per se* […] in bilingual children" [original emphasis] (p. 180), rather attributes their enhanced metalinguistic ability to their "control of attention" (p. 180). The benefit of bilingualism on language learning is also addressed in studies by Ringbom (1987) and Thomas (1988) and will be elucidated in the subsequent chapter on language aptitude. "Ringbom (1987) compared monolingual and bilingual (Finnish – Swedish) learners in Finland learning English as their third language and found that the bilinguals outperformed the monolinguals" (Jessner, 2008, p. 17). Similarly, "Thomas (1988) showed that English-Spanish bilingual students performed significantly better than their monolingual peers when learning French in the classroom" (Jessner, 2008, p. 17). Thus, previous research predominantly assumes that "[b]ilingualism has a positive mediating effect on third language learning" (Cenoz & Valencia, 1994, p. 204).

Arguably, the ability of individuals to acquire languages is closely linked to their language aptitude. According to Carroll, the creator of the Modern Language Aptitude Test (MLAT), language aptitude may be "defined […] as simply an ability or 'knack' for learning foreign languages. Virtually everyone has this ability, but some people appear to learn at a faster rate than others" (Stansfield, 2014). Abrahamsson and Hyltenstam (2008) consider language learning aptitude "as a largely innate, *relatively fixed* talent for learning languages" [my emphasis] (p. 485). Skehan (2002) however, notes that this assumption is not exactly grounded in sufficient research, "but, for now, following Carroll, we will assume that aptitude does not change with the seasons" (p.79). Discussions have also arisen surrounding the question "whether aptitude is largely fixed at birth, or is the result of early experience" (Skehan, 2002, p.79). Skehan (2002) considers research on the "critical period for language learning" (p.82) only to conclude that it has proven to be inconclusive. The majority of previous "research on the role of foreign language learning aptitude in L2 learning has been carried out in classroom contexts" (DeKeyser, 2000, p.506). With regard to *explicit testing*, Krashen (1977) has noted "that aptitude is an important predictor of success in explicit learning […]" (Abrahamsson & Hyltenstam, 2008, p. 486). Contrary to this assumption, Abrahamsson and Hyltenstam (2008) refer to studies by DeKeyser (2000), Harley and Hart (1997), and Robinson (1997), when they write that they all

> have suggested […] that language aptitude may play a decisive role in naturalistic SLA [second language acquisition] – and perhaps an even more decisive role than it plays in instructed SLA – because acquiring a language implicitly, by having to discover grammatical regularities and phonetic patterns merely from language exposure, can be seen as an even greater challenge than learning it through pronunciation tutoring and explicit grammar instruction […]. (p. 486)

Overall, it can be said that previous research has shown to indicate "the centrality of aptitude in relation to L2 acquisition process" (Skehan, 2002, p. 83). Especially memory and L2 learning have been shown to be closely related. Skehan (2002) has noted that "there are higher correlations between the memory component of aptitude and foreign language achievement amongst children with an early age of L2 learning onset […]" (p. 76). Revisiting the initially mentioned assertion that language aptitude can be regarded as being "largely innate" and "relatively fixed" (Abrahamsson & Hyltenstam, 2008, p.485), it ought to be noted that it also "has been found to be relatively independent of other factors, [such as] general intelligence, personality, attitudes toward the language to be learned, and the motivation to learn it" (Abrahamsson & Hyltenstam, 2008, p. 485). Thompson (2013) refers to "Skehan's (1990) explanation that aptitude consists of two separate factors: a person's innate ability for processing language, combined with environmental factors, such as parental language background and literacy" (p. 686). Hence, language aptitude may be considered as being a complex system that can be connected to external factors, albeit not influenced by them. Even so, Sternberg (2002) does believe that "[m]uch of what appears to be foreign-language learning aptitude may reflect a valuing process" (p. 19), which he detected in the disparate opinions regarding foreign language learning in Belgium. While Flemish native speakers feel the need to acquire a second or third language, this is not the case the other way around (Sternberg, 2002,). Sternberg (2002) anticipates that this contrast is closely connected to the assumed practicality of the respective language. Svara (2009) investigates the role of motivation on future language learning and points out that language learning aptitude alone does not result in an actual higher learning curve, but "motivation has a decisive impact on student's willingness and effort invested in language study" (p. 1). Thus, even though language aptitude can be a good predictor for a person's potential to learn a foreign language, to what extent that person makes use of that potential is closely connected to many other factors, such as motivation and language attitudes.

Language attitudes may be understood "as being hypothetical constructs [that] are inferred, conceptual inventions hopefully aiding the description and explanation of behaviour" (Baker, 1988, p. 114) and can be seen as "learned predispositions, [that] are likely to be relatively stable over time" (Baker, 1988, p. 114). Contrary to that, Adegbija (1994) asserts that "attitudes are complex phenomena which could be observable or internal, or both simultaneously, temporary or lasting, and of a surface-level or deep-rooted nature" (pp. 255–256). Duan (2004) simply understands "[l]anguage attitudes to refer to people's feelings and preferences towards their own language and other speech varieties around them" and particularly notes the "value they place on those languages" (p. 12). This idea of valuing a language has already been touched upon above in relation to language aptitudes. Further, it clarifies the subjectivity of language attitudes. Burgaski (1990) has noted the importance of language attitudes as “forces shapping [sic] the intricate dynamics of life in complex human societies” (p. 46), since they reflect group belonging. Similarly, Roos (1990), who conducted a study centring on students’ language attitudes, emphasises this “importance of language as [a] symbol of membership of a certain cultural or ethnic group” (p.27). She further refers to Trudgill (1983) who has also noted the importance of “language [as] a crucial factor that determines group identification and solidarity […]” (Roos, 1990, p. 27).

A possible division within this broad understanding of language attitudes is provided by Fasold (1984), who differentiates between “attitudes towards language itself”, “attitudes towards speakers of a particular language or dialect” and “all sorts of behaviour concerning language […] including attitudes towards language maintenance and planning efforts” (p. 148). This wide-ranging comprehension of language attitudes is also shared by Baker (1992), including “[a]ttitudes to a specific minority language” and “[a]ttitudes to language groups, communities and minorities” (p.29) which both will be of particular relevance in the subsequent study. Adegbija (1994) similarly distinguishes between language attitudes as “accommodate[ing] evaluative judgements made about a language or its variety, its speakers, towards efforts at promoting, maintaining or planning a language, or even towards learning and teaching it” (pp. 255–256). A number of studies have investigated in how far language attitudes may differ between bilingual and monolingual speakers, with a particular focus on bilingual speakers of minority languages. Korani (2012), for instance, conducted a study focusing on bilingual students’ attitudes towards their L2, Farsi, while also enquiring into the diminishing number of speakers of their L1, Kordi, and potential consequences thereof. Korani (2012) particularly noted a gendered as well as residential difference, with females having a more positive attitude towards Kordi, and students living in the city having a more positive attitude towards Farsi. A further study also centring on attitudes of speakers towards minority languages was conducted by “Lambert, Hodgson, Gardner and Fillenbaum (1960) [who] found that French Canadians tend to downgrade their own language group, in contrast to English Canadians, who rate own group more favourable […]” (Ihemere, 2006, p. 196). As opposed to that, “Bourhis, Giles and Tajfel (1973) and Bourhis and Giles (1976) found that Welsh respondents […] judg[ed] speakers of Welsh and of Welsh accented English more favourably than speakers of R.P. English […]” (Ihemere, 2006, p. 196). Adegbija (2000) conducted his own research in West Africa and found that “[s]adly, most of the indigenous languages [in West Africa] are in fact confined in their educational functions to the first three years of primary education” (pp. 82–83). An apparent relation between language attitudes and prestige may be reflected in parents’ decision of teaching a minority language in which they are still proficient to their children. It can be said that if the parents’ attitudes are negative towards a minority language, this will inevitably affect children not only in their attitudes, but also, potentially, in their proficiency in the respective languages. Ihemere (2006) refers to Adegbija (1994) and Batibo (2005) when he writes that is has been observed “that most parents wish their children to have proficiency in the former colonial language to improve their chances of social promotion and economic advancement, even at the expense of their mother tongue" (p. 204). This once again depicts the problematic valuing of one language over another and the key fact of prestige and alleged future benefits of one language when compared to a second. Several scholars have investigated this in relation to future developments of languages. Bell (2013), for instance, in her study on Australian Aborigines focuses in particular on language revival and maintenance and notes that "language maintenance […] is applicable in those situations where there are still reasonable numbers of language speakers [who are] making a concerted effort to keep the language strong […]" (pp. 399–400). This *concerted effort* is seemingly oftentimes lost in the transition of an older generation to a younger. Ihemere (2006) has investigated the shift in attitudes and its relation to age and discovered a "change in language choice, with older speakers using only or mainly Ikwerre dominant patterns in wide ranging contexts and younger speakers utilise both Ikwerre and NPE or principally NPE to fulfil various communicative functions" (p. 205). This shift in languages, which will be discussed in more detail in the following paragraph, may result in younger generations regretting their lack of speaking a certain diminishing language. Bell (2013), for instance, mentions that she feels like an Aboriginal, even though her first language is Australian English, but she and the people in her community "constantly use words and phrases and sing songs in traditional languages that have been passed down through generations" (p. 408). Nevertheless, she points out that "[w]e also feel sadness, regret and sometimes anger that we did not have a chance to speak the 'languages of the land', our heritage languages" (p. 408). However, even though, retrospectively a positive attitude towards a minority language may be expressed, a previous continuous and potentially still increasing negative attitude towards a certain language may have already caused a language shift to occur.

Fasold (1984), for one, provides a categorisation of "social factors causing language shift based on previous studies […]" including "school languages and other government pressures, […] high prestige for the language being shifted to, and small population of the speakers of the language being shifted from" (Duan, 2004, pp. 29–30). However, being an underrepresented minority group with a minority language does not automatically have to result in a shift towards the more prestigious official language of the country. Fase, Jaspaert and Kroon (1992) declare that "[a]s long as there is a minority group, as long as the minority group is not demographically broken up, the use of the minority language will not disappear unless the norms for language use within the groups are changed" (p. 7). Gal (1979) has exactly noted this change of norms, as she refers to a "widespread variation in language use [which] reflected a language shift in progress" (p. xi). Additionally, Gal (1979) has also noted the relation of both language attitudes and language use to age. "She […] found […] that age was a significant factor in predicting use of German rather than Hungarian. Younger people were using German in domains where older people use only Hungarian, which seems to indicate that there is a language shift in progress" (Duan, 2004, p. 24). Yuan (2001) discovered "that the process of a shift from one language to another can be completed in four generations.

She also made the observation that due to centralized education and mass media, the process of language change is accelerating at a faster speed" (Duan, 2004, p. 26). Another reason for the occurrence of language shift is provided by Fishman who "suggests that language shift may occur by an increasing encroachment of one language on the (former) domain of the other [and he] observes that language shift may take place without loss of social identity" (Wald, 1974, p. 314). The culmination of language shift may be language death, which, however frightening it may sound, can nevertheless be considered a thoroughly natural process. Aitchson (1995) "asserts [that] after a language was born and passed its childhood, it grows up and for some reason it will die" (Korani, 2012, p. 1748). Additionally, it is simply stated that "[w]hen people are ashamed of using a language […] it causes language death" (Korani, 2012, p. 1748), which once again emphasises the importance of language attitudes in relation to language use and group belonging. This has also been addressed by Bell (2013) in relation to Aboriginal languages in Australia, who refers to Pensalfini (2003) who noted that one of them, Jingulu, is "in the final stages of its existence" (as cited in Bell, 2013, p. 405), which he links to the lack of older generations teaching the language to the children.

Considering the above stated research on bilingualism, which has predominantly shown its advantageousness for future language learning, and the importance of language attitudes concerning the maintenance of minority languages, the case of the minority group of Burgenland Croats is particularly special. Especially the differences between mono- and bilingual Burgenland Croats with regard to their foreign language aptitude and language attitudes towards Burgenland Croatian are of interest. In order to understand the situation of Burgenland Croats better, it is necessary to introduce some background information at this point: Altogether there are four language groups in Burgenland, namely German, Romani, Hungarian and Croatian. Each federal state in Austria is able to individually determine the administrative status of minority languages (Baumgartner, 2001). The main differences in the handling of minority languages, for instance between Carinthia and Burgenland, are predominantly determined by the respective political party or parties in power (Baumgartner, 2000). As for the Burgenland Croats' historical development, the language group of this minority came to Burgenland in the sixteenth century, in times of the Ottoman Wars (Baumgartner, 2001). Burgenland Croats speak a certain dialect of Croatian, which already shows certain differences within Burgenland, let alone to standard Croatian, and which was standardised to one literary language in the nineteenth century (Baumgartner, 2000). Also in the nineteenth century, it has received support from the government in Vienna, especially in terms of its cultural development (Baumgartner, 2000). The Austrian Independence Treaty from 1955 secured the rights of Burgenland Croats; yet many of those designated rights are still not implemented today (Baumgartner, 2001). In the beginning of the 1960s, linguistic assimilation was increasing in all of Austria, which was not only the case because many believed that they did not speak the corresponding minority language well, or well enough, and therefore favoured German (Baumgartner, 2001), but also due to the political situation in Austria. Nevertheless, over the second half of the twentieth century, many initiatives against assimilation were undertaken, such as a bilingual grammar school, and Croatian TV- and radio programs, as well as Burgenland Croatian being declared the second official language in all but one district in Burgenland in 1987 (Baumgartner, 2000). Bilingual education is an integral part of the educational system of Burgenland and also part of the federal education act ("Landesschulgesetz") (Baumgartner, 2001). A law for minority groups from 1997 defined that classes ought to be held in Burgenland Croatian all over Burgenland, and not only in such communities where minorities are notably present (Baumgartner, 2001). Even though certain rights have been granted to minority languages, it can nevertheless be said that, nowadays, as in previous decades, minority groups and their languages only receive as many rights as necessary (Baumgartner, 2000), as opposed to as many rights as possible.

Monolingual Burgenland Croats have been classified as having solely learned German, with their older relatives still being proficient in Burgenland Croatian, while bilingual Burgenland Croats have acquired both Burgenland Croatian and German at home. Based on the previously described research on bilingualism, H1 expects that bilingual Burgenland Croats have a higher language aptitude than monolingual Burgenland Croats. The second hypothesis is concerned with language attitudes, since it has been shown that they are vital in relation to language use, and potential language shift and death. H2 therefore assumes that bilingual Burgenland Croats have a more positive attitude towards Burgenland Croatian.

2 Methodology

2.1 Participants

The sample of Burgenland Croats of the following cross sectional study consisted of exactly 13 mono- and bilingual Burgenland Croats, resulting in 26 participants in total. As stated above, monolingual Burgenland Croats consider German to be their mother tongue, whereas bilingual Burgenland Croats are proficient in both German and Burgenland Croatian. Altogether, more females than males participated in the study, namely 17 women as opposed to 9 men. There are 9 female bilingual Burgenland Croats and only 4 men, while there are 8 female monolingual Burgenland Croats and 5 men.

2.2 Instruments

In order to test for language aptitude, Part V of the Modern Language Aptitude Test (MLAT), called paired associates, was used. Paired associates consists of 24 word pairs (originally Kurdish – English, but it was adapted to Kurdish – German for the purpose of this study) and each participant has 2 minutes to memorise "and then to choose the meaning of the given word from the multiple choice options provided" (Sasaki, 2012, p. 316). Part V aims to unveil one's "capacity to make associations between L1 lexis and target language items" (Skehan, 2002, p. 71).

To further test for language aptitude, the LLAMA_B was used. The LLAMA_B is part of the LLAMA test battery and "is a simple vocabulary learning task, which measures your ability to learn relatively large amounts of vocabulary in a relatively short space of time" (2005, p. 5). As opposed to the MLAT Part V, however, the participants' memory capacity is tested with the task of memorising 20 images of animal-like figures, also within 2 minutes. "The words to be learned are real words taken from a Central American language, and they are arbitrarily assigned to the target image" (Meara, 2005, p. 5). Crucially, the LLAMA_B, as opposed to the MLAT Part V, is applicable to everybody, irrespective of their L1 (Meara, 2005, p. 5).

In order to test the second hypothesis of the potential difference between mono- and bilingual Burgenland Croats' language attitudes towards Burgenland Croatian, an attitudes questionnaire with ten concepts was compiled, each of which had four to seven questions. The concepts were titled: *doings* (whether the state or local community is doing enough for the maintenance of the minority language), *identity* (in how far do the participants identify as Burgenland Croats), *future general* (how they believe the future of Burgenland Croatian will look like), *family and friends* (in what language they converse/d with each other), *future benefit* (whether they believe that being proficient in Burgenland Croatian has any future benefits for them or their children), *discrimination* (whether they have ever felt discriminated against on the basis of their Burgenland Croatian descent), *everyday doings* (in how far Burgenland Croatian is part of their everyday lives), *personal doings* (how they believe they could improve their competence in Burgenland Croatian), *general competence* (how they would assess the proficiency in Burgenland Croatian of the members of their communities), and, finally, *self-assessment* (as its title indicates, the participants were asked to assess their own proficiency in Burgenland Croatian in very general terms). Aside from this questionnaire, a further general questionnaire was to be filled out, in which basic questions such as age, education and the number of languages participants know, were asked.

2.3 Procedures

The 26 participants were tested over the course of several months and the tests and questionnaires were, in most instances, administered with the presence of the study's conductor. This enabled unrecorded dialog on how the participants perceive their belonging to the minority group of Burgenland Croats, which helped gain a multi-facetted image of their personal impressions and opinions. In select cases, where a meeting proved impossible, the questionnaires and tests were sent via e-mail and with clear instructions. Nevertheless, in all cases the questionnaires were the first items for the participants to fill out, followed by the MLAT Part V and the

LLAMA_B test. This order was primarily compiled due to time reasons, since in several cases a number of Burgenland Croats did the study simultaneously, and starting with the questionnaires enabled everybody to do the MLAT Part V concurrently as well. Since the LLAMA_B online version was used, this occasionally proved problematic in cases where several people were tested at once, since only a limited number of electronic devices were available. For the tests, the participants had to be in a quiet environment, which also at times proved slightly difficult, since some participants distracted one another, especially when attempting to answer the language aptitude tests. Overall it can be said that the participants interacted well with the tests, though most found the LLAMA_B to be considerably more demanding than the MLAT Part V.

3 Results

Data was analysed using SPSS (Version 23). The descriptive parameters show that the MLAT Part V is normally distributed ($p = .068$), whereas the LLAMA_B is not ($p = .002$). A great difference can also be seen concerning the means and minimum and maximum scores of each test. The mean for the MLAT Part V is 12.15 from an attainable 24 points, while the mean is considerably lower with 4.62 out of 20 for the LLAMA_B. The minimum (4) and maximum (24) of the MLAT Part V are much higher than the ones of the LLAMA_B (1 and 11, respectively). The most frequent scores of the MLAT Part V is 7, 11 and 12, with three participants scoring each, whereas the most frequent scores of the LLAMA_B is 1, 3 and 5, which were each attained five times.

As the questionnaire on language attitudes uses Likert scales, Cronbach's α was calculated to test for reliability. After removing some questions in order to improve reliability, the following values were obtained. The 'doings' subscale consisted of 7 items ($\alpha = .780$), 'identity' consisted of 7 as well ($\alpha = .925$) and 'general future' had 5 items (α = .583). The Cronbach's alpha for the 6 'family & friends' items was $\alpha = .879$, while the 5 'future benefit' items had an alpha of .644. The 'discrimination' subscale consisted of 5 items ($\alpha = .534$) and 'everyday doings' of 4 items ($\alpha = .821$). The subscale of 'personal doings' with 6 items proved fairly reliable ($\alpha = .743$), similar to 'general competence' with 6 items ($\alpha = .808$), and, lastly, the subscale 'self-assessment', which consisted of 5 items ($\alpha = .825$).

In order to test whether bilingual Burgenland Croats have a higher language aptitude than monolinguals, a t-test for independent samples was conducted. There was no significant difference found between monolingual Burgenland Croats ($M = 13.69$, $SD = 5.80$) and bilingual Burgenland Croats ($M = 10.69$, $SD = 5.56$); $t(24) = -1.31$, $p = 0.202$.

The independent samples t-test between these groups in the LLAMA_B shows similar results: There was no significant difference found between monolingual Burgenland Croats ($M = 5.08$, $SD = 3.55$) and bilingual Burgenland Croats ($M = 4.15$, $SD = 3.63$), $t(24) = -0.681$, $p = 0.502$. Hence, H1, which assumes that bilingual Burgenland Croats have a higher language aptitude than monolingual Burgenland Croats, has to be rejected.

Contrary to the insignificance of the t-tests for the MLAT Part V and LLAMA_B scores, the t-test for the attitudes questionnaire shows different results. H2 assumed that bilingual Burgenland Croats have a more positive language attitude towards the minority language of Burgenland Croatian than monolinguals. All of the attitudes are normally distributed and the independent samples t-test shows that there is a significant difference in language attitudes towards Burgenland Croatian between monolinguals (M = 6.23, SD = 1.11) and bilinguals (M = 3.51, SD = 1.37); $t(24) = -5.56$), $p < 0.001$. Thus, H2 can be accepted. The closer the participants scored towards 10, the higher were their attitudes towards Burgenland Croatian and its minority group.

4 Discussion

As noted previously, the research question of this paper was concerned with the extent to which language aptitudes and language attitudes may differ between mono- and bilingual Burgenland Croats, which ensued in a H1 that stated that bilingual Burgenland Croats have a higher language aptitude than monolinguals, and a H2 that assumed that bilingual Burgenland Croats have a more positive attitude than their monolingual counterparts.

H1 had to be rejected, since the t-test for independent samples for both the MLAT Part V and the LLAMA_B scores were insignificant. This result clearly contrasts the studies that were discussed in the introduction, which mostly attributed advantages in language aptitude to bi- and multilinguals. Planchon and Ellis (2014) noted in their comparative study of monolinguals' and bilinguals' language aptitudes that the bilinguals had higher results than monolinguals, which they traced back to their increased 'metalinguistic awareness'. Likewise results were obtained by Ringbom (1987) and Thomas (1988) who, in their studies on mono- and bilingual Finish-Swedish and English-Spanish speakers' L3 learning found that bilinguals outperformed monolinguals (Jessner 2008). The results of the language aptitude's t-tests have not only been insignificant, but they have even shown that the mean of the bilingual Burgenland Croats' MLAT scores was slightly lower with 10.69 points, as opposed to 13.62 points for monolinguals. Similarly, the mean for the LLAMA_B scores was higher for the monolinguals with 5.08 points when compared with bilinguals (4.15). Thus, there actually appears to be a slight tendency for monolingual Burgenland Croats, whose native tongue is German, to have performed better. Considering the reviewed literature of the introduction, these results are fairly unexpected.

H2, which assumed that bilingual Burgenland Croats would have a more positive attitude towards Burgenland Croatian, can be accepted since the result of the t-test for independent samples was highly significant ($p < .001$). Considering "the importance of language as [a] symbol of membership of a certain cultural or ethnic group"

(Roos, 1990, p.27), this result does not come as a surprise. The positive language attitudes expressed by bilingual Burgenland Croats indicate their desire for the maintenance of their language group. Bell (2013) has noted this importance of “a concerted effort” (pp. 399–400) in order for languages to be maintained in her study focusing on Aboriginal languages in Australia. She notes the differences in positive and negative, as well as changeable and fixed, language attitudes in relation to the survival and revival of certain Aboriginal languages. The positive attitude of bilingual Burgenland Croats can be seen to indicate the continuing survival of the minority language. Language attitudes are closely connected to evaluating languages and valuing some languages above others (Duan, 2004). This has also been related to language aptitude, in that the negative valuing of a language and the negative language attitude towards a language, may, irrespective of an individual’s language aptitude, influence the speaker’s motivation to acquire a certain language. Thus, it may be deduced from the more negative language attitude expressed by monolingual Burgenland Croats, that it is highly unlikely for them to learn Burgenland Croatian in their remaining lifetime.

Regarding the valuing process that seems intrinsic to language attitudes, Lambert et al. (1960) have shown in their study on French Canadians that bilinguals tend to value English higher than French, whereas, Bourhis et al. (1973), as well as Bourhis and Giles (1976), ascertained that Welsh speakers value their minority language and non-RP speaking individuals higher than standard English (Ihemere, 2006). This study on Burgenland Croats corroborates rather with the study conducted by Bourhis, Giles and Tajfel and Bourhis and Giles, since bilingual Burgenland Croats clearly acknowledge the importance of their minority language, as opposed to the majority of monolingual Burgenland Croats. Further, as indicated by Bell (2013) and Adegbija (2000) in their studies on English in West Africa, the parents’ attitudes towards a language are of great importance for the further maintenance of it. Parents’ negative attitude towards a language may easily result in them not teaching their children the respective language, which may potentially be (one of) their native tongues, thereby fostering language shift, and, as an ultimate consequence, language death. As for their studies on Africa, both Adgebija (1994) and Batibo (2005) have noted “that most parents wish their children to have proficiency in the former colonial language to improve their chances of social promotion and economic advancement, even at the expense of their mother tongue” (as cited in Ihemere, 2006, p. 204). Even though all Burgenland Croats obviously acknowledge the importance of German, as it is the official language of Austria, bilingual Burgenland Croats seemingly do not desire their children to grow up monolingually, but wish to maintain a bilingual upbringing. Thus, even though Aitchson (1995) has remarked that the natural process of one language is to be born and eventually die, this does not seem to be the case for Burgenland Croatian in the near future, judging by the positive attitudes of bilingual Burgenland Croats towards their minority language. Yet, as noted in the historical synopsis of the development of Burgenland Croatian over the course of centuries, a language shift towards German was observable, predominantly due to political circumstances, in the middle of the twentieth century. Overall it can be said that the great influence of various “government pressures”, as well as "high prestige of the language being shifted to and small population of the speakers of the language being shifted from" are significant factors of language shift (Fasold, 1984; as cited in Duan, 2004, pp. 29–30). Nevertheless, judging from the significant t-test of the attitudes questionnaire which has shown that bilingual Burgenland Croats have a more positive attitude towards the minority language than monolingual Burgenland Croats, bilingual Burgenland Croats clearly desire to maintain Burgenland Croatian by continuing to raise their children bilingually.

5 Conclusion

Since the majority of the reviewed literature indicated that bilinguals scored higher on language aptitude tests since bilingualism results in many cognitive advantages, it was assumed that bilingual Burgenland Croats have a higher language aptitude than monolingual Burgenland Croats. However, H1 had to be rejected, since the t-test of independent samples of the administered language aptitude tests MLAT Part V and LLAMA_B, showed no significant differences between mono- and bilingual Burgenland Croats.

Even though this first hypothesis had to be rejected, the second hypothesis, which suggested that bilingual Burgenland Croats have a more positive attitude towards Burgenland Croatian than monolingual Burgenland Croats, yielded significant results. Thus, it can clearly be stated that bilingual Burgenland Croats, those individuals who are proficient in Burgenland Croatian as well as German, think more highly of Burgenland Croatian than monolingual Burgenland Croats, whose sole mother tongue is German. Judging by this positive attitude towards Burgenland Croatian by its speakers, it cannot be clearly stated whether a language shift towards German is in fact in progress or not. In order to determine that, further research would have to be conducted. In case any future investigations on this or a similar topic were to be administered, the number of participants should definitely be increased. Additionally, in order to more adequately test participants' aptitudes, the entire aptitude test batteries ought to be applied, since if one is not extremely successful in the memory part, one may excel in others. Also, as noted in the introduction of this paper, particularly the memory part that was applied in this study has been repeatedly criticised over the years. It would also be productive to test the participants' proficiency in German and Burgenland Croatian, as this may help to classify the groups more adequately. In order to receive a better understanding of a language shift happening, it would be helpful if there were a larger number of younger as well as older participants. In relation to this, a closer investigation of families where the parents are still bilingual Burgenland Croats, in contrast to their children, may provide interesting insights into the micro-level of language shift. By discovering the reasons for deciding not to teach children the language, language shift and death could be closely investigated. This probably is closely linked to language attitudes not only on an individual, but also on a group and society level, which once again may be related to the local and/or national political situation.

Additionally, a comparative study between Burgenland Croats and the other minority languages in Burgenland, such as Hungarian and Romani, may be worthwhile, as well as comparing the situation of Burgenland Croats to Carinthian Slovenes, keeping in mind the differing circumstances in the federal states.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481–509. https://doi.org/10.1017/S027226310808073X

Adegbija, E. (1994). English and indigenous languages in Kwara state (Nigeria): The bottom-line attitudinal factors. *Multilingua – Journal of Cross-Cultural and Interlanguage Communication, 13*(3), 253–284.

Adegbija, E. (2000). Language attitudes in West Africa. *International Journal of the Sociology of Language, 141*(1), 75–100. https://doi.org/10.1515/ijsl.2000.141.75

Aitchison, J. (1995). *Language change: Progress or decay?* Cambridge, UK: Cambridge University Press.

Baker, C. (1988). *Key issues in bilingualism and bilingual education*. Philadelphia: Multilinguial Matters.

Baker, C. (1992). *Attitudes and language*. Clevedon, UK: Multilingual Matters.

Batibo, H. (2005). *Language decline and death in Africa: Causes consequences, and challenges*. Clevedon, UK/Buffalo, NY: Multilingual Matters.

Baumgartner, G. (2000). Sprachgruppen und Mehrsprachigkeit im Burgenland. In H. Van Uffelen, M. Hünning, & U. Vogl (Eds.), *Musik, Sprache, Identität* (pp. 35–73). Vienna: Melzer Kopie.

Baumgartner, G. (2001). Die Minderheitensprache Kroatisch in Österreich und ihre Bildungseinrichtungen. In H. Van Uffelen & C. Van Baalen (Eds.), *Europa der Regionen* (pp. 112–140). Vienna: Melzer Kopie.

Bell, J. (2013). Language attitudes and language revival/survival. *Journal of Multilingual and Multicultural Development, 34*(4), 399–410. https://doi.org/10.1080/01434632.2013.794812

Bialystok, E. (2001). Metalinguistic aspects of bilingual processing. *Annual Review of Applied Linguistics, 21*, 169–181. https://doi.org/10.1017/S0267190501000101

Bourhis, R. Y., & Giles, H. (1976). The language of cooperation in Wales: A field study. *Languaeg Sciences, 42*, 13–16.

Bourhis, R. Y., Giles, H., & Tajfel, H. (1973). Language as a determination of Welsh identity. *European Journal of Social Psychology, 3*(4), 447–460. https://doi.org/10.1002/ejsp.2420030407

Burgaski, R. (1990). The social basis of language conflict and language attitudes. In P. Nelde (Ed.), *Plurilinga: Language attitude and language conflict* (pp. 41–47). Bonn, Germany: Dümmler.

Carroll, J. B. (1973). Implications of aptitude test research and psycholinguistic theory for foreign language teaching. *International Journal of PsychoLinguistics, 2*, 5–14.

Cenoz, J., & Valencia, J. F. (1994). Additive trilingualism: Evidence from the Basque Country. *Applied PsychoLinguistics, 15*(2), 195–207.

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews, 15*(5), 336–343. https://doi.org/10.1038/nrn3709

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*(4), 499–533. https://doi.org/10.13140/RG.2.1.3958.4486

Duan, L. (2004). *A sociolinguistic study of language and language attitudes among the Bai people in Jianchuan county*, China. MA thesis, Payap University, Chiang Mai, Thailand.

Fase, W., Jaspaert, K., & Kroon, S. (1992). *Maintenance and loss of minority languages. Studies in bilingualism*. Amsterdam/Philadelphia: J.Benjamins.
Fasold, R. (1984). *The sociolinguistics of society*. Oxford, UK: Blackwell.
Gal, S. (1979). *Language shift: Social determinants of linguistic change in bilingual Austria*. New York: Academic.
Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communicational effectiveness of English children in French schools. *Canadian Journal of Education, 2*(3), 15–24.
Harley, B., & Hart, D. (1997). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 19*, 379–400.
Ihemere, K. U. (2006). An integrated approach to the study of language attitudes and change in Nigeria: The case of the Ikwerre of Port Harcourt City. In O. F. Arasanyin & M. A. Pemberton (Eds.), *Selected proceedings of the 36th annual conference on African linguistics* (pp. 194–207). Somerville, MA: Cascadilla Proceedings Project.
Jessner, U. (2008). Teaching third languages: Findings, trends and challenges. *Language Teaching, 41*(1), 15–56. https://doi.org/10.1017/S0261444807004739
Korani, A. (2012). Comparison of the attitude of bilingual (Kordi and Farsi) boy and girl students in Islamic Azad University Islamabad Branch toward Farsi as their second language and the effect of the place of their living (city or village) on their attitude. *Procedia – Social and Behavioural Sciences, 47*, 1747–1750. https://doi.org/10.1016/j.sbspro.2012.06.894
Krashen, S. D. (1977). The monitor model of adult second language performance. In M. Burt, H. C. Dulay, & M. B. Finocchiaro (Eds.), *Viewpoints on English as a second language* (pp. 152–161). New York: Regents.
Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology, 60*(1), 44–51.
Liedtke, W. W., & Nelson, L. D. (1968). Concept formation and bilingualism. *Alberta Journal of Educational Research, 14*, 225–232.
Lognostics. http://www.lognostics.co.uk/tools/llama/. Accessed 29 Mar 2016.
Meara, P. (2005). *LLAMA language aptitude test: The manual*. Swansea, UK: University of Wales.
Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs, 76*(27), 1–23.
Pensalfini, R. (2003). *A grammar of Jingulu, an aboriginal language of the northern territory*. Canberra, Australia: Pacific Linguistics.
Planchon, A., & Ellis, E. (2014). A diplomatic advantage? The effects of bilingualism and formal language training on language aptitude amongst Australian diplomatic officers. *Language Awareness, 23*(3), 203–219.
Ringbom, H. (1987). *The role of L1 in foreign language learning*. Clevedon, UK: Multilingual Matters.
Robinson, P. (1997). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning, 47*, 45–99.
Roos, R. (1990). Language attitudes in the second language situation. *Per Linguam: A Journal of Language Learning, 6*(2), 25–30.
Sasaki, M. (2012). The modern language aptitude test (paper-and-pencil version). *Language Testing, 29*(2), 315–321.
Skehan, P. (1990). The relationship between native and foreign language ability: Educational and linguistic factors. In H. Dechert (Ed.), *Current trends in European second language acquisition research* (pp. 83–106). Clevedon, UK: Multilingual Matters.
Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–93). Philadelphia/Amsterdam: John Benjamins.
Stansfield, C. (2014). What is language aptitude? *Language learning and testing foundation.* http://lltf.net/aptitude-tests/what-is-language-aptitude/. Accessed 29 Mar 2016.
Sternberg, R. J. (2002). The theory of successful intelligence and its implications for language-aptitude testing. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 13–43). Philadelphia/Amsterdam: John Benjamins.

Svara, C. (2009). *English reconsidered: Investigating language attitudes of Austrian high school and vocational college students*. Diploma thesis, University of Vienna.
Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual and Multicultural Development, 9*, 235–246.
Thompson, A. S. (2013). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *The Modern Language Journal, 97*(3), 658–701.
Trudgill, P. (1983). *Sociolinguistics: An introduction to language and society*. Middlesex, ON: Penguin.
Wald, B. (1974). Bilingualism. *Annual Review of Anthropology, 31*(1), 301–321.
Weinreich, U. (1966). *Languages in contact: Findings and problems*. The Hague, The Netherlands: Mouton.
Yuan, Y. (2001). *Language contact and language change: A case study of Atsang*. Beijing, China: Beijing Minority Press.

The Correlation of Early Multilingualism and Language Aptitude

Sofia Hörder

Abstract Since the 1960s, studies comparing bilinguals and multilinguals to monolinguals have found bi/multilinguals to be advantaged in several linguistic areas, including language aptitude. Linguists believe this to be due to multilinguals exhibiting, among other things, a heightened metalinguistic awareness. Since most studies, however, rarely differentiate between bi/multilingualism acquired before the age of five and bi/multilingualism language competence acquired later in life, this study set out to investigate the interface of *early* bi/multilingualism and language aptitude, hypothesizing that *early* bi/multilinguals would outperform *late* bi/multilinguals on language aptitude tests. We analyzed the test scores of thirty-seven 15–35 year-old participants based in Vienna, who sat the MLAT III, MLAT IV, MLAT IV tests, as well as the LLAMA_B. They were categorized as monolinguals (n = 11), bilinguals (n = 17) or multilinguals (n = 9) according to the age of onset of their languages. Attitudinal and motivational factors were additionally investigated with the help of a questionnaire and then correlated with the test scores. As a result, this study found the overall test scores of early bi/multilinguals to be statistically indistinguishable from monolinguals (*late* bi/multilinguals), meaning *early* bi/multilinguals were *not* found to score better at language aptitude tests than monolinguals. When analyzing the sub-scores, it was surprisingly only in terms of *grammatical sensitivity* (as measured by the MLAT IV) that *mono*linguals significantly (p = .036) outperformed bilinguals and multilinguals. Factors such as *gender*, *education, motivation, facility of acquisition* and *time invested into improving language skills* were also found to have a significant impact on the participant's language aptitude scores.

S. Hörder (✉)
Department of English and American Studies, University of Vienna, Vienna, Austria
e-mail: a1301213@unet.univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_15

1 Introduction

There has been a substantial amount of research done on the factors that influence an individual's language aptitude, yet very few studies have examined whether or not being raised bilingually or even multilingually has a positive effect on language aptitude. It is an important question to ask in a globalized world, where being raised as a monolingual is becoming increasingly rare, except perhaps in countries such as the USA, UK, Australia and China where one language is still dominant. Our intuition seems to suggest that logically the more languages someone knows and learns early on, the easier it will be for them to acquire additional languages. This is because linguists assume that bi/multilinguals have more linguistic knowledge and tools at their disposal, which they can transfer from one language to another. The topic is, however, a problematic one, because it doesn't seem 'fair' that people who are raised with multiple languages should then additionally be privileged in terms of language aptitude, since being raised in a particular manner is pure chance and cannot be influenced by the individuals themselves.

The most specific studies on this topic so far are Thompson's *The Interface of Language Aptitude and Multilingualism: Reconsidering the Bilingual/Multilingual Dichotomy* (2013), which studied the relationship between language aptitude and previous language experience in Brazilian language learners and Planchon and Ellis's *A diplomatic advantage? The effects of bilingualism and formal language training on language aptitude amongst Australian diplomatic officers* (2012), which focused on whether bilingualism has an effect on one's ability to learn subsequent languages. Both these studies found positive results, concluding that there is a significant positive correlation between being raised with more than one language *or* having 'early' foreign language experience or training, and the ability to learn further languages quicker and with more facility. Both studies, however, focus less on the aspect of *being raised* with more than one language and its influence on language aptitude and rather on *any previous experience* with foreign languages (i.e. in school or professional training) and language aptitude.

1.1 Monolingualism, Bilingualism and Multilingualism

As Jessner (2006, p. 1) points out, "[m]ultilingualism is a growing phenomenon and certainly not an aberration […] but a normal necessity for the world's majority". Despite multilingualism being in fact 'the norm', establishing a working definition of multilingualism or even of bilingualism is considerably challenging. Baker (2011, p. 15) holds that being able to define a bilingual is "elusive and ultimately impossible". There are nevertheless instances where a definition is necessary, as for this present study. The difficulty with defining both bilingualism and multilingualism stems from a long history of discussion concerning this phenomenon. The central aspect of debate having been the *fractional* versus the *holistic* view of multilingualism. In an article from 1985, Grosjean (1985, pp. 467–77) begins to argue against what is now called the monolingual view of multilingualism. This view holds that "multilingual speaker[s] are seen as several monolinguals in one person" (fractional) (Jessner, 2006, p. 10), as opposed to "perfectly competent speaker-hearers in their own right" (holistic) (Grosjean, 2008, p. 12f.). This definition holds that multilingual speakers have a native-like control of each of their respective languages. After studying bi/multilinguals for years, Grosjean came to the conclusion that the communicative competence of bilinguals did *not* equate to being "the sum of two complete or incomplete monolinguals, but [rather] a unique linguistic profile" (Baker & Prys Jones, 1998, p. 9). More modern definitions now consider this type of bilingualism – called 'balanced bilingualism' – as only *one* possible kind of bilingualism along a *continuum*. Bilingualism and multilingualism are now made dependent of factors such as "competence, cognitive organization, age of acquisition, exogeneity, socio-cultural status, and cultural identity" (Hamers & Blanc, 2000, p. 49). Baker (2011, p. 15) and Mackey (1970, pp. 554–584) both strongly highlight the communicative aspect of language and differentiate between language *use* and *ability*. Making active use of a language is therefore considered essential in being classified as a bi−/multilingual, rather than simply amassing theoretical knowledge about a language that is never put into practice. This study also does not hold to ideas about a so called 'critical period' for language learning, which sets an age limit on people being able to become bilingual (usually between 21 months and puberty). As Baker and Prys Jones (1998) state, "a critical period of language development is now discredited", but "there are advantageous periods" (p. 660). Therefore, this study assumes that anyone can become bilingual given the circumstances, but is interested in whether language acquired early in this 'advantageous period' (i.e. before the age of 5) impacts an individual's language aptitude differently than language competence acquired later on.

In their study of Australian bilinguals, Planchon and Ellis (2012, p. 204) made use of Hamers and Blanc's (2000) definition of bilingualism (or as they term it: '*bilinguality*') as "the psychological state of an individual who has access to more than one linguistic code as a means of social communication" (p. 25). In the following, this paper will also use this definition of bilingualism, as it focuses on the ability to actively communicate in a language. This paper will, however, add to the definition the requirement that bilinguals could communicate in more than one language before the age of five. The definition also applies to multilingualism by referring to having 'access to *more than one* linguistic code' [original emphasis], yet the term 'bilingual' will also be differentiated from 'multilingual', since the distinction is relevant for the analysis. The term *bilingual* is therefore ascribed to individuals who can access *two* languages and the term *multilingual* as pertaining to individuals who have access to *three or more* languages. Less progressively, this paper does categorize individuals as monolinguals, bilinguals or multilinguals according to the age of acquisition/onset of their language. Even if a participant actively uses more than one language (even native-speaker-like), but only acquired their second language after the age of five (usually at the age of 8/9 for a second language or 11/12 for a third language) they will be classified as monolinguals in this study. This is due to the fact, that this study is particularly interested in the impact of languages acquired at an *early* age (before learning a foreign language at school) on language aptitude and whether the influence of this type of bi−/multilingualism can in fact be differentiated from other types of subsequent language learning in this context at all.

1.2 Multilingualism and Third or Subsequent Language Acquisition

It is now increasingly recognized by various linguists that bilingualism plays a positive role in facilitating third language acquisition (TLA) (Bialystok, 2009; Cenoz, 2003a, 2003b; Grosjean, 1985; Thompson, 2013). This is due to a variety of reasons. According to Jessner (2006), the research shows that the abilities that speakers, who are competent in more than one language, gain, result in "higher levels of metalinguistic awareness, creativity or divergent thinking, communicative sensitivity and the facilitation of additional language acquisition by exploiting the cognitive and linguistic mechanisms underlying these processes of transfer and enhancement" (p. 27). TLA therefore strongly differs from second language acquisition (SLA) "as a result of the prior linguistic knowledge of the bilingual or multilingual language learner" and consequently merits "a research focus in its own right" (Planchon & Ellis, 2012, p. 205).

1.2.1 Metalinguistic Awareness and Mental Flexibility

In the following, this paper will examine these claims further. The attribute which is probably most commonly ascribed to bi-/multilinguals in linguistics is that of exhibiting a higher level of metalinguistic awareness when compared to monolinguals. Jessner (2006) defines metalinguistic awareness as: "the ability to focus attention on language as an object in itself or to think abstractly about language and, consequently, to play with or manipulate language" (p. 42). Researchers first ascribed this type of cognitive advantage to bilinguals when the Peal and Lambert study was published in 1962. In their study, Peal and Lambert (1962) found that bilinguals exhibited "mental flexibility, a superiority in concept formation [and] a more diversified set of mental abilities" (p. 20). As a result of his study of bilinguals and monolinguals in the Basque country, Lasagabaster (1997) also found that bilinguals exhibited a higher level of metalinguistic awareness than monolinguals. These findings were then increasingly supported by other linguists (Bialystock, 2009; Micheal & Gollan, 2009; De Groot, 2011). Additionally, linguists such as Bialystock (2009), have found that "bilingualism endowed children with enhanced mental flexibility, and this flexibility was evident across all domains of thought" (p. 418). In other words, linguists found mental flexibility to be an attribute of the multilingual mind. Mental flexibility is connected with metalinguistic awareness, as mental flexibility is required to become aware of how one is manipulating language and managing to successfully switch between languages. One trait of mental flexibility is the ability to suppress interference from other languages in one's repertoire when speaking one language, while simultaneously making use of cross-linguistic connections and concepts from the other languages to enrich the way multilinguals communicate. Michael and Gollan (2009) give an example of how bilingualism results in mental flexibility; Bi−/Multilinguals constantly experience "interference from competing [languages]" (p. 395), since their languages are continually activated simultaneously. By being constantly confronted with this challenge, multilinguals unconsciously focus less on areas like "semantic structure and concreteness" as they have "language specific retrieval cues", which in turn allows them to be less restricted by a language's rules and structure and more flexible in their thinking about language (Michael & Gollan, 2009, p. 399). Thus Bee Chin Ng and Wigglesworth (2007) have concluded that "bilinguals have an advantage when it comes to analyzing language forms, owing to their earlier exposure to two different linguistic codes, since such exposure promotes a more analytic orientation to linguistic operations" (p. 62).

1.2.2 Mechanisms of Transfer

Concerning the aspect of cross-linguistic transfer, bi/multilinguals are said to have an advantage in this area (Jessner, 2006, p. 42). As Jarvis and Pavlenko (2008) found, multilinguals are "open to make more interlingual identifications between the target language and previously learned languages" (p. 205) than monolinguals. They concluded that "[t]he more languages people know, the more likely they are to exhibit transfer from one or more of those languages" (Jarvis & Pavlenko, 2008, p. 205). More specifically, multilinguals display less inhibitions concerning the structure and especially the "grammaticality of the target language" (Jarvis & Pavlenko, 2008, p. 205), because they have access to a larger and more varied linguistic repertoire and are thus more likely to experiment and transfer their existing skills to the target language. This is advantageous for language learning (and consequently language aptitude) in that it represents yet another tool- namely that of transfer- that multilinguals may employ when learning new languages. In Cenoz' (2003a) words: "Monolinguals tend to formulate grammars that are just powerful enough to fit the input data, that is their grammars are more restricted but include fewer errors. Multilinguals generate larger grammars which include incorrect sentences but allow them to progress faster" (p. 79). Transfer thus plays a particularly positive role in terms of communicative aspects of language learning. And although it might result in more mistakes made in the foreign language, it ultimately contributes to a faster pace of learning by means of trial and error.

Another more specific aspect of multilingual transfer has been examined by Kemp (2007), who set out to test whether multilinguals outperform bilinguals in terms of foreign grammar acquisition. Her final results showed that the amount of languages an individual knew related to the amount of grammar learning strategies they made use of when learning a foreign language (Kemp, 2007, pp. 256–57). Additionally, multilinguals were confirmed to have a higher aptitude for foreign grammar learning due to automatizing their previously acquired strategies and transferring them to the foreign language being learned (Kemp, 2007, pp. 256–57). This study will therefore be especially interested in the results of the MLAT IV in regards to bilingual status.

1.2.3 Communicative Sensitivity

According to Jessner (2006, p. 27), among others, bi/multilinguals exhibit a more distinctly developed level of communicative sensitivity than monolinguals. Alcón Soler (2012) found evidence for this in one of her studies, finding that "[b]ilinguals display[ed] a higher degree of pragmatic awareness and higher communicative sensitivity" (p. 530) and that "bilinguals show evidence of higher communicative sensitivity, mainly in the form of concern for their interlocutors' feelings, and follow a hearer-oriented communicative approach" (Alcón Soler, 2012, p. 530). This characteristic allows bi/multilinguals to interact with their interlocutors more successfully, which, consequently, enables them to learn and adapt to the pragmatics that are entailed by the foreign language's culture, more rapidly. Language aptitude tests are often criticized for focusing too much on formal aspects of language learning and being less appropriate in gauging an individual's informal and communicative success in language acquisition (see section on aptitude testing below). Since this factor is not considered in the test battery and therefore does not influence the results of this study, but may, nevertheless, be an important factor of reality, the author of this study found it important to include this aspect in reflecting on bi/multilinguals in relation to language aptitude.

1.2.4 Cases of Monolingual Superiority in Regards to Third Language Acquisition (TLA)

The majority of studies agree that bi/multilinguals outperform monolinguals on language aptitude tests (Cenoz, 2003a, 2003b; Kemp, 2007; Planchon & Ellis, 2012; Thompson, 2013). There are, however, also studies which found opposite results that also need to be acknowledged. Although they are far smaller in number and tend to be viewed as the exception, these studies found either *no* difference between monolinguals and bilinguals (Schoonen et al., 2002) or cases where monolinguals delivered superior results compared to bilinguals in relation to TLA. Van Gelderen et al. (2003), for example, compared reading comprehension scores of monolingual Dutch or bilingual Dutch (and other language) participants in L2 (for monolinguals)/L3 (for bilinguals) English reading competence. They found that "the L3 readers performed significantly lower" and concluded that "the advantage of being bilingual (in spoken language) may not pay off for reading and writing development in the L2 and a third foreign language" (Van Gelderen et al., 2003, p. 22). They did provide several reasons for why this might have been the case, including that they did not control for the participant's educational or socioeconomic background which might have skewed the results. Moreover, they figured that “the L1 – L3 gap for [their] bilingual group [might have been] bigger than the L1 – L2 gap for the monolinguals” (Van Gelderen et al., 2003, p. 22), which would make it “possible that an advantage in metacognitive strategy use for L3 readers was not sufficient to overcome disadvantages at the lower level of reading” (Van Gelderen et al., 2003, p. 22). Several other studies which found such results (although not all) also faced similar difficulties according to Cenoz (2003a, p. 76). Yet even Cenoz, who found many cases of bi/multilingual superiority in terms of metalinguistic awareness and general competence in regards to TLA, also found cases where bilinguals obtained lower results than monolinguals (Cenoz, 2003b, p. 114f.).

All in all, according to research done so far, bilinguals and multilinguals exhibit enhanced metalinguistic awareness and mental flexibility, as well as a number of other qualities (e.g. superior concept formation and communicative sensitivity), which benefit their foreign language acquisition rate and thereby enhances their language aptitude. The minority of studies found results where monolinguals outperform bi/multilinguals in language aptitude tests. When studies find such results, they tend to explain them by pointing to other interfering factors such as sample size or socio-economic-background.

1.3 Language Aptitude Testing

Language aptitude as defined by Dörnyei (2005) is “a number of cognitive factors making up a composite measure that can be referred to as the learner’s overall capacity to master a foreign language” (pp. 33–34). Skehan’s (1990) definition states that language aptitude consists of a combination of “a person’s innate ability for processing language”, “as well as environmental factors such as literacy and parental language background” (pp. 83–85). Thompson (2013, p. 686) explains that researchers were first interested in measuring language aptitude in American schools in the 1920s, in order to identify and invest in linguistically gifted students. This approach was then taken up by the American military to increase the cost-effectiveness of their language training programs (Dörnyei, 2005, p. 33ff.). A considerable amount of language tests have in fact been developed by the United States Department of Defense, including the Defense Language Aptitude Battery (DLAB) which Planchon and Ellis (2012) used in their study. Carroll (1981) cautions, however, not to confound *aptitude* and *achievement*. While *aptitude* refers to a person’s “current state of capability”, *achievement* reflects an individual’s “actual performance” (Carroll, 1981, p. 84). Factors such as motivation or even general intelligence, for example, may play a far larger role than aptitude in terms of *achievement* (Cenoz, 2003a, p. 74). Moreover, aptitude “does not place limits on learning a language, but it appears to affect the speed of language learning in more formal language [learning settings]” (Baker & Prys Jones, 1998, p. 656). So the extent to which an individual’s aptitude as measured by aptitude tests is indicative of their actual language learning success may depend on various other factors as well.

1.4 Purpose and Hypothesis

Taking these and other studies into account, this study aims to contribute to the literature on effects of multilingualism, as well as on third or subsequent language acquisition, in regard to their relation to language aptitude. The **hypothesis** therefore put forth and tested in this paper is that persons raised with two or more languages from birth or before the age of five, will outperform those who gained language learning experience later in life on language aptitude tests. Bilinguals and multilinguals, are therefore expected to outperform those who have had previous language experience (i.e. in school or professional training), but were not raised with multiple languages from before the age of 5. Based on the research done in similar fields so far, we assume this to be due to the fact that the bi-and multilinguals possess a more developed level of metalinguistic awareness (Jessner, 2006, p. 42). More specifically, bi/multilinguals are thought to have at their disposal a higher grammatical sensibility, inductive language ability (Planchon & Ellis, 2012, pp. 204–206) as well as cognitive flexibility concerning things such as higher tolerance of ambiguity (Bialystock, 2009, pp. 417–25). They also employ more tools and strategies when learning a new language in terms of more divergent and creative thinking (Paradis, 2008, p. 414), associative memory, use of mnemonics and transferring elements from their other languages (Jarvis & Pavlenko, 2008; Kemp, 2007). The **null-hypothesis** to be disproven is therefore as follows: There is no difference in the language aptitude scores of early bi/multilinguals and monolinguals (i.e. late bi/multilinguals). However, as Cenoz (2003a, p. 74) points out, it is also important to consider the possibility that bilingualism may not affect all aspects of third language proficiency. This paper will therefore not only take the overall score of the MLAT and LLAMA_B into consideration, but will also examine the sub scores of the individual parts as they relate to the bi/multilinguals' language aptitude. In a second step, factors such as gender, education, motivation, facility of acquisition and time invested into improving language skills will also be controlled for, to see whether they have any significant impact on the participant's language aptitude scores aside from the age of acquisition. This will provide a more complete picture concerning the main hypothesis that aims to be answered, as these factors may have a greater impact than the age of second or third language acquisition or, conversely, none at all.

2 Methods

2.1 Participants

This study tested a sample of 37 participants on their language aptitude. The subject group was obtained from a larger pool of students [only two of whom were language students (of a language other than English)]. This ensured random selection

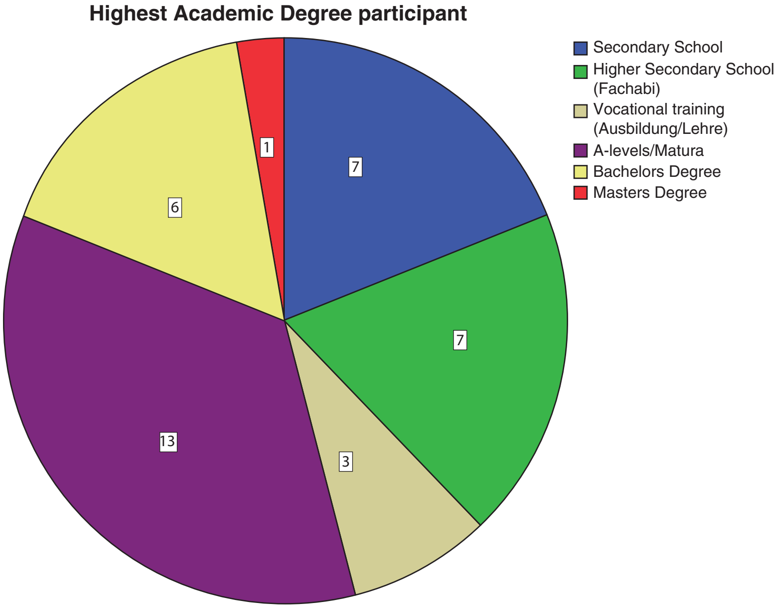


Fig. 1 Participant's highest academic degree

and a rich variety of multilinguals who spoke a combined number of 121 languages (amount of total languages spoken by participants combined).

Albeit coming from differing language backgrounds, 27 out of 37 of the participants had similar educational qualifications (Higher Secondary School or A-levels (completed) and pursuing A-levels or BA degree) Fig. 1 and were close in age (81.1% of participants were between the ages of 17 and 25 (with 18% being between 25 and 35 and 1 fifteen-year-old) see Fig. 2.

They also all currently live in a German speaking country and are exposed to German. The majority is also exposed to English, be it at home, in school or at church and with friends. It is only in their third and subsequent languages that the participants (especially the multilinguals) differed greatly, as they spoke a variety of different languages from Korean to Kwi. The ratio of male (43.2%) to female (56.7%) participants was almost balanced as well.

Monolinguals made up 29.7% of the testing group, Bilinguals 45.9% and Multilinguals 24.3% (Bilinguals and Multilinguals combined = 70.2%) see Fig. 3. All participants spoke two or more languages but the monolinguals acquired these subsequent languages after the age of five.

The cut-off age of five has been selected based on Hyltenstam's and Montrul's findings in combination with more pragmatic reasons. As a result of his study of Swedish bilinguals, half of which acquired their second language before the age of 6 and half after the age of 7, Hyltenstam (1992) has shown that loss of linguistic ability may set in as early as age 6, which is why age five was chosen as the age prior to this potential hampering of linguistic ability. Montrul adds that "[e]xisting

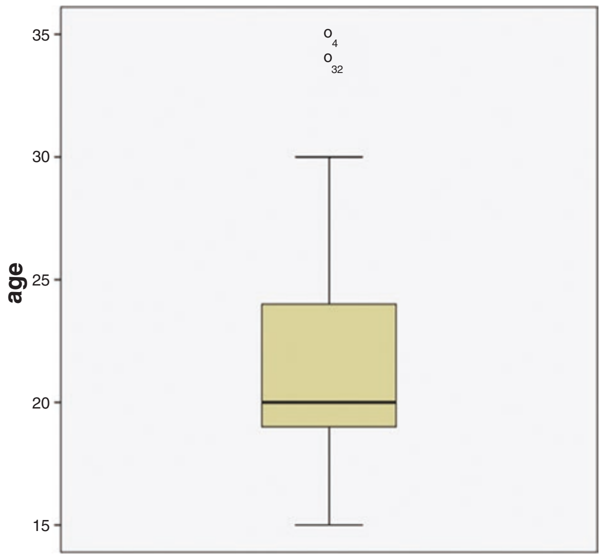


Fig. 2 Participant's age

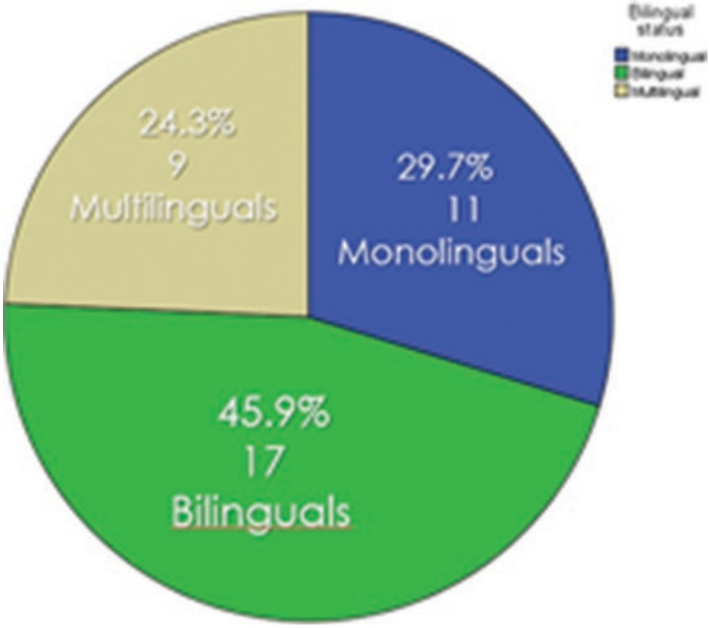


Fig. 3 Participant's bilingual status

case studies point to the conclusion that, if linguistic input and socialization occur before the ages of 6–7, chances of developing a full linguistic system are remarkably good" (2008, p. 16). Combined with Montrul's finding that "[e]arly child L2 acquisition probably spans about two years and occurs between the ages of 4–6, when spoken language is practically fully developed but the children have not yet received formal schooling" (2008, p.17), the given information lead to the selection of the compromise cutoff age of 5 years. This age also coincides with the school entry age of most countries, where additional languages are presumably studied, meaning they are explicitly learned, rather than acquired through the child's environment as is the case with a native language. Furthermore, it is still a quite early age, despite it not being as early as 3 years old, and since it was difficult to find enough bilinguals under the age of three the scope was broadened to the "compromise age" of five, which is still supported by Montrul's findings.

2.2 Instruments

Developed by Carroll and Sapon in 1959, the Modern Language Aptitude Test (MLAT) series is one of the most commonly used tests for measuring language aptitude. It consists of five parts: "1. Number learning, 2. Phonetic Script, 3. Spelling Clues, 4. Words in Sentences, 5. Paired Associates" and has been "validated only for students who know English with a native or native like fluency" (Tiedeman, 1960, pp. 582–84). In his review of the MLAT, Tiedeman allows that "the last three parts may be used as a short form of the test" (1960, pp. 582–84). This is presumably due to time constraints, as the full version of the test takes up to 70 min (Tiedeman, 1960). Even though the MLAT has become less popular over time, because as Thompson (2013) points out "researchers feel that the subtests of the MLAT measure aptitude as it relates to a more audiolingual pedagogical framework" (p. 686), the "test has proven to be relevant even in the most current teaching and research contexts" according to Stansfield and Reed (2004, p. 44). The MLAT is still considered to be adequate to predict success at the initial stages of foreign language learning, no matter the teaching approaches (Ehrman, 1998). The fourth part of the test (words in sentences) is so far the only one 'exempt from criticism' as DeKeyser (2000) states that this part "is specifically aimed at measuring grammatical sensitivity and therefore should be the best predictor of grammar learning" (p. 509).

The LLAMA_B test was developed by Paul Michael Meara and the Vocabulary Acquisition Research Group at the Swansea University in 2005 and is a vocabulary learning task comparable to the MLAT V, yet different in some key aspects. As in the MLAT V, participants are also given 2 min to study vocabulary items and are then tested on their ability to assign them correctly. The difference lies in the fact that the LLAMA_B is a 'language free' test. Made-up words are assigned to picture stimuli and these are then mixed before the participant has to assign the correct word to the correct picture. This study is also interested in seeing how the MLAT V (with English->Kurdish words) and the LLAMA_B (made-up words->pictures) correlate.

The MLAT was therefore used for this study as a useful predictor of foreign language learning success. It has also been found to have similar predictions as subsequently developed tests such as the DLAB and the Pimsleur Language Aptitude Battery (PLAB) according to Ehrman (1994, p. 75) and is the most easily accessible out of the three. Ehrman (1994) in fact found the MLAT to be superior to other predictors concerning language learning achievements (p. 90). The LLAMA_B test was also chosen as an easily accessible and administrable test in order to examine whether it correlated with the MLAT V and supported the results.

2.3 Procedures

The participants were given a questionnaire[1] in addition to the language aptitude test battery. They were asked to provide their language competence history and give more specific details for every language they knew (were able to communicate in) such as the age of onset, their proficiency (on a scale of 1–10 (1 = minimum, 10 = native like command of the language), and marking their languages as their L1 or L2 or L3 etc. (in order of acquisition). Their language proficiency scores (from 1–10) were combined to create a Combined Language Competence Score (CLCS). Therefore, a person who speaks English (proficiency 10), German (9), French (7) and Arabic (5) received a CLCS of 31. Whereas a person who speaks English (10) and Spanish (5) received a CLCS of 15. This was done in order to be able to analyze whether being *raised* as a bi/multilingual was distinguishable from simply having a lot of previous language experience that was acquired after the age of 5 (i.e. a monolingual with a high CLCS).

Participants were also asked to specify how they were raised (from birth or before the age of 5), with several options being provided for monolingualism/bilingualism/trilingualism and multilingualism respectively (or 'other'). These were two of the options for being raised bilingually:

1. My mother and father spoke different languages and one of these languages was predominant in the environment.
2. My mother and father spoke the same language, but there was a big influence from another language from my environment (i.e. country, daycare, nanny etc.).

Participants were additionally asked to provide general background information such as their age, nationality, gender, place of birth, place of current residence and mobility background. Another piece of information that was elicited, was the participant's academic background (with their highest completed degree), as well as the academic background of their parents. This was done in order to identify any external factors that render the group more heterogeneous which could have skewed or influenced the results. The final part of the questionnaire consisted of seven questions aimed at evaluating the participants' attitude and motivation towards language

[1] Questionnaire available on request.

learning, their perceived facility of language learning (i.e. self-efficacy), as well as their personal opinion of whether or not they believe that their previously obtained languages facilitate their learning of novel languages. The last question probed for the amount of time participants invested in improving their language skills. This information was elicited with outliers in mind (monolingual in the positive sense; multilinguals in the negative sense), in order to potentially be able to correlate them with a particularly high/low motivation towards language learning. The questions were summed up to achieve a total score for each category (i.e. question a + c = motivation; b + f = facility; d + e = opinion; g = time).

Following the questionnaire the participants took the MLAT III, IV and V tests as well as the LLAMA_B online test, which took the majority between 45 and 55 minutes to complete. Approximately half of the participants took the test in the administrator's presence and the other half took it 'at home'. In the MLAT III the participant's ability to associate sounds with symbols is tested. A word is represented by a selection of letters that suggest the sound of the word. Additionally, there are five words provided, one of which is a type of synonym of the disguised word. The participant must then for example choose the word *affection* for the spelling clue *luv* (love). The MLAT IV tests grammatical sensitivity. Here the participant is given a sentence with a capitalized word and a following sentence with several underlined words. The participant must choose the option which in essence fulfills the same grammatical function as the capitalized word in the first sentence. The MLAT V tests vocabulary learning and speed of association. Here participants are given 2 min to study 24 Kurdish words with their English translations. They must then see how many words they memorized correctly by ticking the correct translation of the words which are switched in order on a separate sheet. The LLAMA_B also tests vocabulary learning and speed of association in a language free context.

The tests and questionnaires were on rare occasions handed back incomplete, but over the course of 4 weeks, the administrator of the test was able to collect the missing data from most of the participant's resulting in a final dropout rate of 18%. The only piece of data that remained incomplete was the information concerning the highest academic degree of the parents. Since this information was not directly relevant for the research question and of a comparatively sensitive nature, the study went ahead and used these participants' data without the parents' education variable.

3 Results

The data collected from the tests and questionnaires were then transferred into SPSS. Once the variables were coded, the study first tested the MLAT and LLAMA_B scores for an underlying normal distribution and homogeneity of variance between groups. The MLAT IV ($p = .200$, *mean* = 21.27, SD = 7.5) and the MLAT V ($p = .200$, *mean* = 16.05, SD = 5.6) were found to be normally distributed according to the Kolmogorov-Smirnov-test, as they are not significant and therefore normally distributed. Although the MLAT III ($p = .005$, *mean* = 36.2, SD = 8.9) and

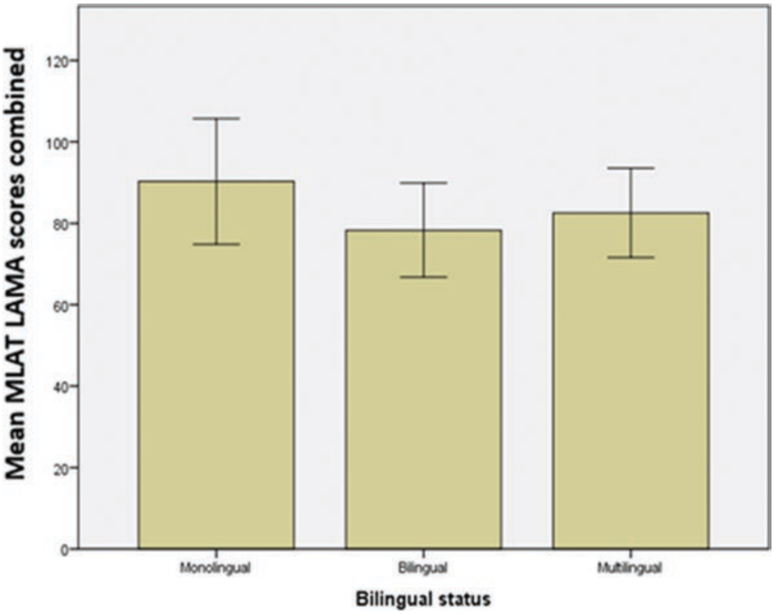


Fig. 4 Correlation of bilingual status and combined scores

LLAMA_B (p = .011, *mean* = 9.35, SD = 4.3) scores just marginally failed to meet the criterion of normality according to the Kolmogorov-Smirnov-test, they were visually inspected with Q-Q diagrams and found to be within acceptable bounds of normality. According to the LLAMA manual, the average score for the LLAMA_B is 8 and this study's participants were close to that with a score of m = 9.35. Then the study proceeded to test whether being monolingual or bi-/multilingual had an influence on the test scores, independent from other variables. This was followed by examining the influence of other factors on the test scores including English native speaker status, gender, combined language competence score, education, number of countries of residence, number of languages spoken and age. Lastly, the attitudinal factors elicited in the questionnaire such as the participants' motivation, facility of language learning, opinion on the subject, and the amount of time invested in improving language skills, were tested in correlation with the test scores. The effect of monolingualism or bi−/multilingualism was overall statistically insignificant (p = .376).

For the most part, the study yielded unexpected results. The hypothesis was that multilinguals would outperform bilinguals and bilinguals would outperform monolinguals on the MLAT and LLAMA_B tests. The null-hypothesis to be disproven therefore is: There is no significant difference in the language aptitude scores of monolinguals, bilinguals and multilinguals. This null-hypothesis was, however, *confirmed* rather than disproven by the results of this test, at least concerning the overall scores. Contrary to expectations, a bivariate correlation showed, that there was *no* significant correlation between the overall testing scores and bi/multilingual status (Fig. 4).

Table 1 Significance of correlation between bilingual status and combined scores

		Bilingual status	MLAT LAMA scores combined
Bilingual status	Pearson Correlation	1	−.150
	Sig. (2-tailed)		$\boxed{.376}$
	Sum of Squares and Cross-products	19.892	−84.216
	Covariance	.553	−2.339
	N	37	37
MLAT LAMA scores combined	Pearson Correlation	−.150	1
	Sig. (2-tailed)	.376	
	Sum of Squares and Cross-products	−84.216	15903.568
	Covariance	−2.339	441.766
	N	37	37

As can be seen in Table 1, the probability level p = 0.376 is over the significance threshold of 0.05. These results fail to support the bulk of literature in this field that finds bi−/multilingual superiority in most cases concerning language aptitude. It supplies supportive evidence for the contrary and smaller literature of cases which have found either no significant difference between the groups or cases of monolingual superiority. When combining bilinguals and multilinguals as one category and measuring them against the monolinguals, the results remained similar and statistically insignificant.

To further test whether there is no difference in test scores based on the amount of languages spoken, the MLAT IV was chosen as an example for the MLAT tests used in this study. Two independent samples t-tests were conducted: The differences in scores in the MLAT IV between monolinguals (M = 25.73, SD = 7.88) and bilinguals (M = 19.59, SD = 7.83) was insignificant; t(26) = 2.02; p = 0.054. A similarly insignificant result was obtained when looking for differences between bilinguals (M = 19.59, SD = 7.83) and multilinguals (M = 19, SD = 3.67); t(24) = 0.21, p = 0.833. Hence, it can be concluded that there is no benefit in speaking more languages.

Additionally, to check the results that there is no difference between monolinguals and bilinguals, and bilinguals and multilinguals in general, two independent samples t-tests were conducted between these groups for the combined test scores of the LLAMA_B and MLAT tests. Results show that there is a significant difference between monolinguals (M = 90, SD = 17) and bilinguals (M = 79, SD = 10); t(26) = 2.16, p = 0.039. However, there is no difference between bilinguals (M = 79, SD = 10) and multilinguals (M = 82, SD = 9); t(24) = 0.75, p = 0.459. This shows that there is an advantage of monolinguals in the aptitude tests, which is reflective of the correlation analysis that is described in the following paragraph.

When examining the individual sub-scores of the MLAT there are, however, noteworthy significances. Firstly the participants' MLAT sub-scores (Table 2) highly correlate with each other p = 0.000, meaning that whoever performed strongly on one of the MLAT tests also performed strongly on the other two. These results indicate that the results are quite reliable, as they reflect positively on the fact

Table 2 Correlation of bilingual status and sub-scores

		Bilingual status	MLAT III score	MLAT IV score	MLAT V score	LAMAb score
Bilingual status	Pearson Correlation	1	.018	−.346*	−.059	−.089
	Sig. (2-tailed)		.914	.036	.728	.600
	N	37	37	37	37	37
MLAT III score	Pearson Correlation	.018	1	.583**	.671**	.175
	Sig. (2-tailed)	.914		.000	.000	.299
	N	37	37	37	37	37
MLAT IV score	Pearson Correlation	−.346*	.598**	1	.583**	.277
	Sig. (2-tailed)	.036	.000		.000	.097
	N	37	37	37	37	37
MLAT V store	Pearson Correlation	−.059	.671**	.583**	1	.517**
	Sig. (2-tailed)	.728	.000	.000		.001
	N	37	37	37	37	37
LAMAb score	Pearson Correlation	−.089	.175	.277	.517**	1
	Sig. (2-tailed)	.600	.299	.299	.001	
	N	37	37	37	37	37

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)

that those who tested well did so consistently. Moreover, the MLAT proves to yield consistent results concerning an individual's language aptitude, meaning that the MLAT IV and V do not yield contradictory results (i.e. a very positive or a very negative score) for the same participant.

The second correlation concerns the LLAMA_B and the MLAT V scores. As can be seen in Table 2, there is a high significance of $p = 0.001$. A scatterplot yielded the result $r = 0.516$. This result is also encouraging and not unexpected, since they both test the participant's vocabulary learning skills and speed of association. In both tests the participants are given 2 min to learn new vocabulary, the only difference being that Kurdish words are tested in the MLAT V and remembering invented names for cartoon images are the object of the LLAMA_B. It is therefore logical that these scores should correlate and it speaks for the results of this study that they do, yet it is nevertheless interesting to note that this type of learning and speed of association applies to both words and images.

Additionally, and more importantly for the research question of this paper, Table 2 shows a significant correlation between bilingual status and the MLAT IV $p = 0.036$. Surprisingly, the scores correlate significantly with those of the *mono*linguals (i.e. *late* bi/multilinguals) (see Fig. 5).

The overall mean for the MLAT IV was 21, 27, the monolinguals achieved a mean of m = 25, 73 (SD = 7, 88) the bilinguals a mean of m = 19, 59 (SD = 7, 83)

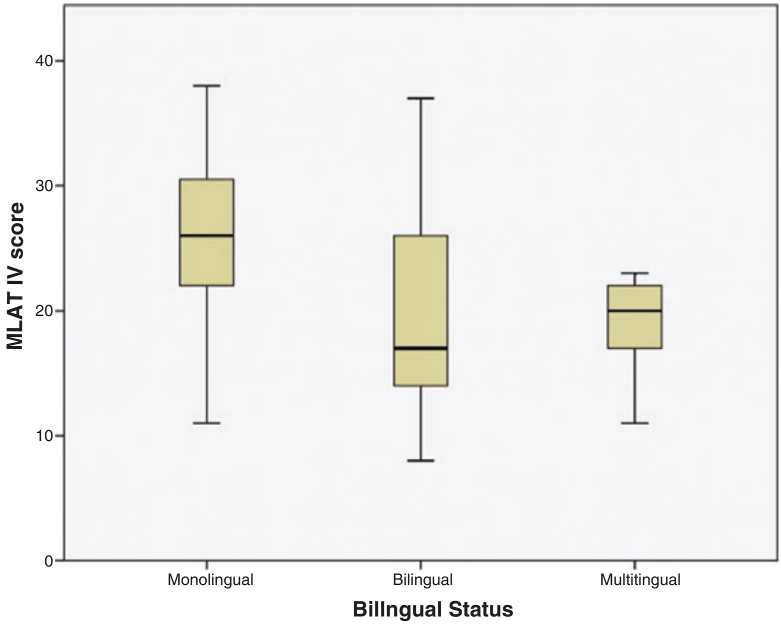


Fig. 5 Boxplot of MLAT IV mean scores and bilingual status

and the multilinguals a mean of m = 19, 00 (SD = 3, 67). As previously mentioned, the MLAT IV is so far the only one exempt from criticism as DeKeyser (2000) states that this part “is specifically aimed at measuring grammatical sensitivity and therefore should be the best predictor of grammar learning” (p. 509). It is therefore interesting to note that it is monolinguals who performed better in this category.

Moving back to the main hypothesis, concerning the overall scores. No significant difference was found in the scores of mono-/bi and multilinguals in the overall scores, meaning that the hypothesis that persons raised with two or more languages from birth or before the age of five, will outperform those who gained language learning experience later in life on language aptitude tests was not confirmed. Therefore this following section will consider further factors such as *gender*, *education, motivation, facility of acquisition* and *time invested into improving language skills* to see if they had a more significant impact on language aptitude test scores than the age of second or third language acquisition.

3.1 Scores and English Native Speaker Status

Since the MLAT is not a language free test and required a high proficiency in English, it was important to control for a potential correlation between the test scores and English native speaker status (ENSS). The participants were grouped into a binary system, being categorized as either ‘English native speaker’ or ‘non-native speaker’. Those participants whose L1 is English (their predominant language) were categorized as English native speakers. There was, however, no correlation. The statistical threshold for the MLAT III and ENSS ($p = .426$), MLAT IV and ENSS ($p = .282$), MLAT V and ENSS ($p = .818$) and LLAMA_B and ENSS ($p = .637$), were all not significant.

3.2 Scores and Gender

The results did show a correlation between the MLAT V and LAMA_B scores and gender. Females outperformed males in every subcategory of the MLAT and the LLAMA_B yet the only significant correlations were found for the MLAT V ($p = .008$) and the LAMA_B ($p = .025$). Out of the 24 possible points, the mean score achieved by females in the MLAT V was 18.14 while the mean for the males was 13.31 with a standard deviation of $SD = \pm 5.1$. As the MLAT IV and the LLAMA_B scores correlated before, it makes sense that if females performed better than males on one of these tests, they would also do so for the other one. Although this is a noteworthy and significant result, as gender and language aptitude are not the object of this study, it will not be discussed in further detail in the ensuing discussion. Let it just be said, that these results are in line with the bulk of studies done so far on gender and language aptitude, which have also found females to excel at vocabulary learning tasks and outperform males in this area.

3.3 Scores and Combined Language Competence Score (CLCS)

Originally the idea for the CLCS was to be able to provide an explanation for outliers who might have been raised as monolinguals, but have acquired several additional languages since then. Since the overall test scores did not correlate with bi−/multilingualism, the study controlled for a possible correlation between the participants’ CLCS and the overall and subtest scores. These, however, did not show any significant correlation either. The statistical threshold for the MLAT III and CLCS ($p = .716$), MLAT IV and CLCS ($p = .935$), MLAT V and CLCS ($p = .152$) and the MLAT and LLAMA_B scores combined and CLCS ($p = .509$), were all not significant.

3.4 Scores and Age

Since the age range did span from 15 to 35 (even though 81.1% of participants were between 17 and 25), it was also important to control for a possible correlation between the test scores and the participant’s age. With a probability level of $p = .082$

for the MLAT III and age, p = .106 for the MLAT IV and age, p = .664 for the MLAT V and age and p = .661 for the LLAMA_B score and age, however, age proved to be a non-significant factor in correlation with the test scores.

3.5 Countries of Residence and Languages Spoken

In the questionnaire the participants were also asked for the number of countries they lived in, since some of them (either as diplomats' kids or missionaries' kids) moved around the world quite frequently. In the analysis the correlation between the number of countries someone has lived in (and the different languages they were therefore exposed to) was correlated with the test results. Nevertheless, this yielded no significant correlation as the probability level for the MLAT and LLAMA_B scores combined and the number of countries of residence (p = .956) was below the significant threshold. In addition, the study analyzed if the number of languages someone spoke correlated with the test scores. But again, there was no statistically significant correlation (p = .547).

3.6 Scores and Education

Figure 6 shows how the participant's degree of education correlates with the MLAT III and IV scores. The significant threshold is very high for both the MLAT III (p = .001) and MLAT IV (p = .004). This shows that the higher the participant's academic degree, the better they scored on the MLAT III and IV. The correlation was not significant for the MLAT V (p = .056) nor the LLAMA_B (p = .118). Neither did the academic degree of the mother or the father significantly influence the participant's scores.

3.7 Questionnaire: Motivation, Facility, Opinion and Time

In the questionnaire the participants were asked to answer questions about their motivation toward language learning, about how easy they felt they could acquire new languages, whether in their opinion their aptitude would improve if they knew more languages or not and how much time they spent on improving their language skills, with a Likert scale. Their answers were then correlated with the MLAT and LAMA_B sub scores and yielded intriguing results. The motivation, facility, opinion and time variables all correlated with the MLAT V and the LLAMA_B scores. Motivation and MLAT V (p = .013), facility and MLAT V (p = .029), opinion and MLAT V (p = .038) and time and MLAT V (p = .035). As well as; motivation and LLAMA_B (p = .026), facility and LLAMA_B (p = .015), opinion and LLAMA_B

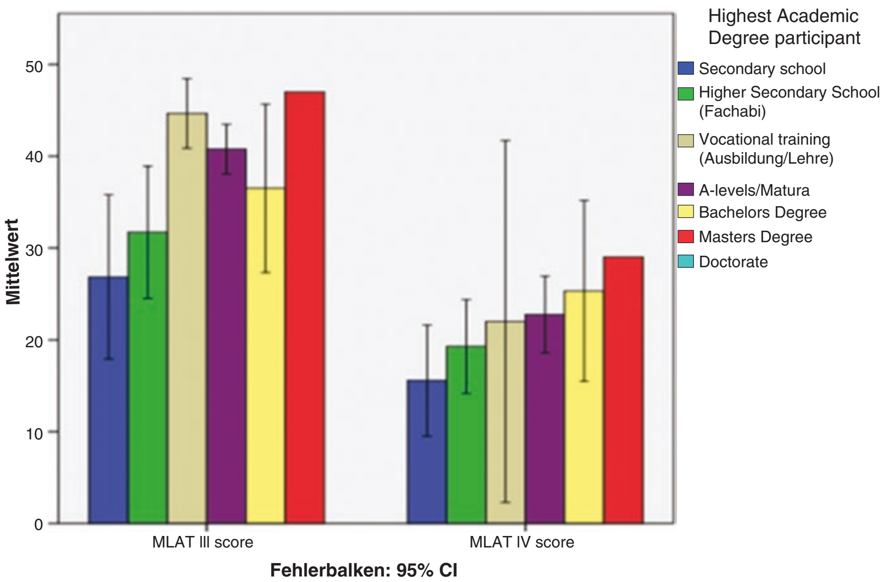


Fig. 6 Bar chart of MLAT III & IV scores and education

(p = .222) and time and LLAMA_B (p = .006). Again, it was to be expected that the MLAT V and LLAMA_B would both yield significant correlations if one of them did. It is interesting to note, however, that these attitudinal and motivational factors only seem to influence the vocabulary learning tests and not the grammar aspects or the participant's ability to associate sounds with symbols.

4 Discussion

Initially, the fact that language aptitude only provides an indication of the *rate* languages are learned at and not the participant's actual *achievements* needs to be recalled. The results presented above, stand contrary to the initial hypothesis. Rather than indicating a positive correlation between bi−/multilingualism and the MLAT and LLAMA_B scores, the only positive correlation found concerning the research question, was between being raised as a *mono*lingual and the MLAT IV scores. These results are inconsistent with the majority of studies conducted in this field which propose that bilingualism has a positive effect on facilitating third language acquisition (Bialystok, 2009; Cenoz, 2003a, 2003b; Grosjean, 1985; Thompson, 2013) due to bi/multilingual's assumed advantages in metalinguistic awareness (Bialystock, 2009; De Groot, 2011; Micheal & Gollan, 2009) and access to a wider range of mechanisms of transfer (Jarvis & Pavlenko, 2008). Nevertheless, these findings contribute to the smaller yet significant body of research which has found such opposite results (Cenoz, 2003a, 2003b; Van Gelderen et al., 2003). Readers must also recall that the participants in this study were categorized into monolinguals, bilinguals and multilinguals according to whether they were raised with one, two or more than two languages from birth or before the age of five. As a result, this study posed a different question than many other studies in this field, which assess the participants' current language capability in their languages or aptitude and not whether they were raised that way or not. It is also important to note that there are consistencies with linguists such as Baker and Prys Jones (1998, p. 655), who state that factors such as attitudes and motivation were more likely to influence language learning success than language aptitude as measured by tests, and Cenoz (2003a), who adds that

> bilingualism was found to exert a significant influence on different measures of English language proficiency such as listening, writing, speaking, reading, grammar and vocabulary. Nevertheless, the effect of factors such as general intelligence and motivation was more important than the influence of bilingualism. (Cenoz, 2003a, p. 75)

As the results showed, factors elicited in the questionnaire concerning motivation, facility, opinion and time correlated significantly with the MLAT V and the LLAMA_B which both tested an essential ability required for language learning, namely that of memorizing vocabulary. Since an IQ test was not conducted for this study, it was not possible to control for the variable of *intelligence*. As Dörnyei (2005) points out, aspects of intelligence share 'definite commonalities' (p. 47) with those of language aptitude, and it therefore appears important to test for this variable in future studies. This study was, however, able to control for *education* and provided significant results for the MLAT III and IV, which are more characteristic of classroom testing situations than the MLAT V and the LLAMA_B. Furthermore, Baker and Prys Jones (1998) point out that "[l]anguage aptitude tests reflect the linguistic rather than the communicative aspects of language learning" (p. 656), meaning that monolinguals (as categorized in this study) who receive instruction of the linguistic, formal structure of a foreign language in school might in fact score higher on language aptitude tests than bilinguals. The bilinguals may, nevertheless, have a higher language aptitude concerning the 'communicative aspects of language learning' (Baker & Prys Jones, 1998, p. 656). As mentioned above in the section on cases of monolingual superiority over multilinguals, Van Gelderen et al. (2003) hypothesize that "the advantage of being bilingual (in spoken language) may not pay off for reading and writing development in the L2 and a third foreign language" (p. 22). This might therefore be an indication for such a case, where a bilingual's ability to communicate in a foreign language does not translate into them possessing a higher language aptitude concerning grammar and the structural aspects of language. Developing a test which measures the communicative and informal aspects of language aptitude as opposed to the formal and structural elements of language might therefore be an imperative direction for future research.

4.1 Monolinguals and Grammatical Sensitivity

This study's result of monolingualism correlating significantly with the MLAT IV but none of the other tests, appears to be problematic at first, since one may generally assume that if one group (i.e. in this case monolinguals) scores higher for one of the tests compared to the other groups, they should do so on the other tests as well. As Cenoz (2003a) points out, however, "even if bilingualism has an effect on third language acquisition, it does not have to affect all aspects of third language proficiency in the same way" (p. 74). In this case we have seen that the participant's bilingual status only has an effect on the aspect of grammatical sensitivity. As mentioned above, Kemp (2007, p. 250) found multilinguals to make use of more grammar learning strategies than monolinguals due to their diversified language learning experience. She also found bilinguals to be significantly outperformed by multilinguals in terms of grammar learning. Kemp (2007) concludes that "[t]he end result of bilinguals' lack of experience may be that working memory is taken up with coping with the cognitive load, and less attention is available for focusing on and internalizing the grammatical form of input" (p. 257), whereas multilinguals have automatized this process. Kemp's results would therefore lead us to assume that multilinguals should have performed better than monolinguals in the MLAT IV. I'd like to put forth two suggestions why this might not have been the case. First of all, the monolinguals in this study were in truth also bi- or multilinguals (after age 5), so the scores/variables relevant for this question are the CLCS and the *number of languages spoken*. When correlated with the MLAT IV scores, however, both variables yielded insignificant results, meaning this study only found 'no difference between the number of languages a participant knew and grammatical sensitivity' as opposed to 'monolinguals outperform multilinguals in grammar learning aptitude' which would contradict Kemp's results. Therefore the results of this study cannot be directly compared with Kemp's due to the differing categorization of monolinguals and multilinguals.

This result is not surprising since it has been noted from experience that multilinguals are known to have difficulties learning grammar (or its formal structure and rules) due to the fact that they often feel like they never 'studied/learned' a language formally, but rather picked it up in informal learning situations while growing up. As for languages they do study, multilinguals tend to rely much more on instinct than grammar rules. This concurs with Cenoz' (2003a, p. 79) findings (see Sect 1.2.2 above) that multilinguals tend to make more grammar mistakes, but these actually help them learn faster. Many native speakers of a language will say that they could not explain their language's grammar rules, but that rather, they know what is correct and what is not instinctively. An influencing factor for the results in this study could, however, be that the monolinguals that participated are not 'pure' monolinguals, since they were only raised in one language, but have studied another or several additional languages, often having a high level of command in these languages. Nevertheless, this experience of *formally studying* grammar in both native and foreign languages- and this naturally translating into a positive score on a formal academic test such as with the MLAT IV- is, in my estimation, the decisive factor for this result.

4.2 Multilinguals and Vocabulary Learning

Arguably, the most surprising result is that multilinguals did not perform better in the vocabulary learning tasks (MLAT V and LLAMA_B). As De Angelis (2007) states that "[b]ilinguals were generally found to perform better than monolinguals on […] vocabulary tests" (p. 119). Concerning one specific study, De Angelis (2007, p. 126) reports the result of Nation and McLaughlin's study in 1986, where multilinguals outperformed monolinguals in a vocabulary learning task but only when the learning was implicit. So it is only when multilinguals can freely choose a vocabulary learning method that they perform better. This can be connected with the assumption that multilinguals tend to use strategies such as mnemonics to remember words more easily. Even though no formal interviews were conducted, we asked our participants what strategy they used to remember the Kurdish words or the images. Those who performed best, stated that they associated the word or image with something related and memorable that they knew (i.e. they used mnemonics). It was, however, not only multilinguals who used these mnemonics, but those who were interested in, or spent a lot of time learning languages. Interestingly, as De Groot and Van Hell (2005) found, "keyword mnemonics are relatively ineffective in experienced FL learners" (p. 13) and more effective in learners with no (or less) foreign language experience. This finding correlates with the results of this study, since aspects of motivation and time both had a significant impact on the participant's score. This again concurs with Cenoz (2003a), who holds that "factors such as general intelligence and motivation [are] more important than the influence of bilingualism" (p. 75).

4.3 Attitudes and Achievement

As already evident in the results, the study found the participants' motivation to learn a language, their perceived facility, time investment, as well as their opinion regarding the usefulness of knowing more languages in regards to learning additional languages, to positively influence their language aptitude scores. This comes as an encouragement to those who complain that bi/multilinguals might have an unfair advantage in regards to language aptitude. But as can be taken from these and other studies (e.g. Cenoz, 2003a, 2003b), monolinguals or persons with less foreign language learning experience, can compensate the advantage bi/multilingualism might provide (according to most studies, e.g. Planchon & Ellis, 2012; Thompson, 2013) in second language learning with motivation, practice and effort. Interestingly, this advantage might be largest in terms of vocabulary acquisition and consequently aspects of memory. Attitudes towards language learning can therefore also be confirmed by this study to have an important impact on language learning in general and vocabulary learning in particular.

4.4 Education or 'Biliteracy' As an Influential Factor

As the results showed a significant correlation between the participants' level of education and the MLAT III and MLAT IV results (i.e. sound-symbol association ability and their grammatical sensitivity), this study examined literature that further elaborated on such findings. De Angelis (2007), for example, references a number of studies which found that it is not bilingualism per se but "bilingual literacy [which] has a crucial role in bringing about positive effects in third language acquisition" (p. 118). Sanz (2000) also explicitly states that it is not bilingualism, but bilingual literacy or 'biliteracy' that is decisive (pp. 23–24). Cenoz also seconds this view (2003a, p. 83). This study (as many others) failed to control for literacy in each of the participant's languages. This is problematic because bilinguals may be individuals who can speak, read and write fluently in two languages, but individuals who can speak a second language fluently and have never learned to read or write in that language might also consider themselves bilingual (as they might for example communicate with a parent in that language). De Angelis (2007) confirms this assumption when reporting the results of another study where "those who had received some formal training outperformed those who had never received any training and learned the language in the home environment." (p. 119). This might be a strongly influential factor which needs to be considered in further studies on the topic.

5 Conclusion

5.1 Limitations and Challenges

An intriguing detail that the data collected in this study showed, was the great diversity of types of multilingualism. To define and categorize individuals as bilinguals or multilinguals posed quite a challenging task due to the fact that there were so many different versions of multilinguals. A challenge other researchers will face as well when conducting such studies, because at the moment we wish to examine multilingualism, we are confronted with multilinguals, who are inherently a strongly diverse and inhomogeneous group. This is due to the fact, that as soon as more than one language is introduced to an individual from birth, this means an introduction to a different culture or cultures, or even countries, which makes it difficult to control for all the potentially invalidating/interfering variables when comparing such individuals to monolinguals. To elaborate upon this, the following cases may be examined:

Case 1: An Austrian participant with parents from the Philippines, who has lived all her life in Austria, but is exposed to Tagalog and English at home.

Case 2: A half Mexican, half Austrian participant who has not lived anywhere longer than 3 ½ years and learned English at an early age in the USA and now speaks English with her siblings, German with the majority of people in Austria and Spanish with her father.

Both participants provided a proficiency level in English of 9. One of them speaks GA English with several people on a daily basis (Case 2), the other mostly with her parents and occasionally with friends and has not lived in an English speaking country, but does consider it a mother tongue (Case 1). The participant in case 2 speaks Spanish only with her father (a similar situation as in case 1 with English) and provides a proficiency of only 6. The trouble with eliciting proficiency scores from participants themselves seems obvious. The problem is exacerbated by others who were raised with two languages and have almost completely lost one, or by those who provide every language they know, even if it is only a small amount, and provide a proficiency of 1 or 2, and those who do not even count that type of language knowledge even though both cases have a comparable proficiency in the language. The variety of English that is spoken differs as well. Researchers naturally should not discriminate between varieties of a language but the problem lies within the test battery, because in my experience the MLAT does discriminate. It was not designed for speakers of Nigerian or Australian English for example who might be exposed to very different vocabulary than English speakers in the USA or UK.

One solution to this problem would be to locate a more homogenous group of participants. However, the flaw of this approach seems to be that is does not reflect the complex reality of the various forms of multilingualism. I therefore dare to say that aiming to differentiate and categorize multilinguals for these types of tests as homogenously as possible, seems a misguided approach if the aim is to reflect the diverse reality of any type of multilingualism. The best proposition for researchers encountering similar difficulties would be to always provide a proficiency test for each of the languages of the participant. Additionally one could potentially create aptitude tests that are sensitive to varieties of English.

Finally, one also has to consider the fact that the monolinguals in this study were mostly (but not entirely) English native speakers, since English was required to partake in the MLAT tests. Even though no significant correlation was found between the ENSS and the scores, monolinguals did outperform bi−/multilinguals in the MLAT V and did better overall (though not significantly) on the MLAT tests. The Monolinguals also tended to be older and therefore have a higher educational degree and as the results showed education did have an impact on the MLAT III and IV scores.

5.2 Directions for Future Research

Deliberating on the outcome of this study, it is clear that further research needs to be conducted to deliver more precise information on the influence of early bi/multilingualism on an individual's language aptitude. As most studies do not differentiate between early bi/multilingualism and bi/multilingualism acquired later in life, more studies on the influence of early bi/multilingualism as opposed to language learning experience acquired later in life on language aptitude should be conducted, in order to provide us with a more complete picture on bi/multilingualism's influence on language aptitude. The following will provide some suggestions on how to tackle this task:

Firstly, more homogeneous groups of participants are needed to partake in such studies. Future studies should, however, focus on a wider range of different age groups and a variety of cultural or linguistic backgrounds and steer away from predominantly investigating children in English or Spanish speaking cultures (as this is what most studies have done).

One essential piece of information that linguists should elicit in future studies is the participant's biliteracy status, and not just the bilingual status, as biliteracy (Cenoz, 2003a, 2003b; De Angelis, 2007; Sanz, 2000) is presumably a much more accurate indicator of bilingual proficiency as it relates to language aptitude. This goes hand in hand with conducting a language proficiency test in each of the participant's languages. Additionally, an IQ test should be administered since Dörnyei (2005) maintains that intelligence and language aptitude are "composite structures, subsuming a number of distinct components" (pp. 45–46) which tend to overlap, in order to control for another potentially significant variable against the aptitude test scores.

It would also be interesting to see a test of verbal or communicative language aptitude developed and for its results to be measured against more formal and analytic language aptitude tests such as the MLAT. This would allow researchers to explore whether there is a difference in the ability of bi/multilinguals in terms of communicative aptitude versus analytic, formal or structural language aptitude.

In conclusion, the results of this study did not disprove the null-hypothesis and did not concur with the majority of studies, meaning that it found there was *no* significant difference in the overall test scores of monolinguals, bilinguals and multilinguals, categorized as such according to the age of acquisition of their languages. The language aptitude scores of early bi/multilinguals was therefore seen to be indistinguishable from those who became bi/multilinguals later in life. The participants' number of languages (i.e. the way multilingualism is generally categorized), however, also showed no significant correlations with the test scores. Meaning, this study also found no correlation between bi/multilingualism in the common sense and language aptitude test scores. It did, however, lack a control group of 'pure' monolinguals. It was only in terms of grammatical sensitivity that monolinguals (i.e. late bi/multilinguals) outperformed multilinguals on the MLAT IV. Other factors which were found to influence the results were: (**1**) Gender: Females outperformed males in the vocabulary aptitude tests; (**2**) Attitudes: Motivation, facility and the amount of time invested correlated with the participants' vocabulary aptitude scores; and (**3**) Education: The higher the participants' educational degree, the better they scored on the MLAT III and IV. The other factors this study controlled for, such as English native speaker status, combined language competence score, age, number of languages spoken, number of countries lived in and the parent's education did not correlate significantly with the test scores. This study does acknowledge, however, that the sample of participants were not as homogenous as presumably required for a study of this character and that this factor most likely influenced the outcome of these results to some extent.

References

Alcón Soler, E. (2012). Teachability and bilingualism: Effects on third language learner's pragmatic knowledge. *Intercultural Pragmatics, 9*(4), 511–541. https://doi.org/10.1515/ip-2012-0028

Baker, C. (2011). *Foundations of bilingual education and bilingualism* (p. 15). Clevedon, UK: Multilingual Matters.

Baker, C., & Prys Jones, S. (1998). *Encyclopedia of bilingualism and bilingual education* (pp. 655–660). Clevedon, UK: Multilingual Matters.

Bialystock, E. (2009). Consequences of bilingualism for cognitive development. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 417–432). Oxford, UK: Oxford University Press.

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 83–118). Rowley, MA: Newbury House.

Cenoz, J. (2003a). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism, 7*(1), 71–89. https://doi.org/10.1177/13670069030070010501

Cenoz, J. (2003b). The role of typology in the organization of the multilingual lexicon. In J. Cenoz, B. Hufeisener, & U. Jessner (Eds.), *The multilingual lexicon* (pp. 103–116). Dordrecht, The Netherland: Kluwer.

De Angelis, G. (2007). *Third or additional language acquisition* (pp. 115–129). Clevedon, UK: Multilingual Matters.

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals* (pp. 385–389). New York: Psychology Press.

De Groot, A. M. B., & van Hell, J. G. (2005). The learning of foreign language vocabulary. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9–29). Oxford, UK/New York: Oxford University Press.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*(4), 499–533.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition* (pp. 33–47). London/New York: Routledge.

Ehrman, M. (1994). A study of the Modern Language Aptitude Test for predicting learning success and advising students. *Language Aptitude Invitational Symposium Program Proceedings* (pp. 75–90). Arlington, VA. Retrieved May 24, 2016, from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SeachValue_0=ED461258&ERICExtSearch_SearchType_0=no&accno=ED461258

Ehrman, M. (1998). The modern language aptitude test for predicting learning success and advising students. *Applied Language Learning, 9*, 31–70.

Grosjean, F. (1985). The bilingual as a competent but specific speaker hearer. *Journal of Multilingual and Multicultural Development, 6*(6), 467–477.

Grosjean, F. (2008). *Studying bilinguals* (p. 12). New York: Oxford University Press.

Hamers, J. F., & Blanc, Michel H. A. (2000 [1989]). *Bilinguality and Bilingualism*. [2nd edition]. Cambridge, UK: Cambridge University Press, 25.

Hyltenstam, K. (1992). Non-native features of near-native speakers: On ultimate attainment of childhood L2 learners. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 351–368). Amsterdam: Elsevier.

Jarvis, S., & Pavlenko, A. (2008). *Cross-linguistic influence in language and cognition* (pp. 203–207). New York: Routledge.

Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language* (pp. 1–42). Edinburgh, UK: Edinburgh University Press.

Kemp, C. (2007). Strategic processing in grammar learning: Do multilinguals use more strategies? *International Journal of Multilingualism, 4*(4), 241–261. https://doi.org/10.2167/ijm099.0

Lasagabaster, D. (1997). *Creatividad y conciencia metalingüística: Incidencia en el aprendizaje del inglés como L3*. [Metalinguistic creativity and consciousness: Impact of learning Elglish as an L3]. Leioa: Universidad del País Vasco.

Mackey, W. F. (1970). The description of bilingualism. In J. Fishman (Ed.), *Readings in the sociology of language* (pp. 554–584). The Hague, The Netherlands: Mouton de Gruyter.

Michael, E. B., & Gollan, T. H. (2009). Being and becoming bilingual. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–399). Oxford, UK: Oxford University Press.

Montrul, S. (2008). *Incomplete Acquisition in Bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins Publishing.

Ng, B. C., & Wigglesworth, G. (2007). *Bilingualism: An advanced research book* (p. 62). London/New York: Routledge.

Paradis, M. (2008). Language and communication disorders in multilinguals. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of the neuroscience of language* (pp. 341–348). London: Academic.

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied, 76*(27), 1–23.

Planchon, A., & Ellis, E. (2012). A diplomatic advantage? The effects of bilingualism and formal language training on language aptitude amongst Australian diplomatic officers. *Language Awareness, 23*(3), 203–219. https://doi.org/10.1080/09658416.2012.742907

Sanz, C. (2000). Bilingual education enhances third language acquisition: Evidence from Catalonia. *Applied PsychoLinguistics, 21*(1), 23–44. https://doi.org/10.1017/S0142716400001028

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Snellings, P., Simis, A., & Stevenson, M. (2002). Linguistic knowledge, metacognitive knowledge and retrieval speed in L1, L2 and EFL writing. In S. Ramsdell & M. Barbier (Eds.), *New directions for research in L2 writing* (pp. 101–122). Dordrecht, The Netherland: Kluwer.

Skehan, P. (1990). The relationship between native and foreign language ability: Educational and linguistic factors. In H. W. Dechert (Ed.), *Current trends in European second language* (pp. 83–106). Clevedon, UK: Multilingual Matters.

Stansfield, C. W., & Reed, D. J. (2004). The story behind the modern language aptitude test: An interview with John B. Carroll (1916-2003). *Language Assessment Quarterly, 1*(1), 43–56. https://doi.org/10.1207/s15434311laq0101_4

Thompson, A. S. (2013). The Interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *The Modern Language Journal, 97*(3), 685–701. https://doi.org/10.1111/j.1540-4781.2013.12034.x

Tiedeman, D. V. (Ed.), & Manuel, H. T. (1960). Testing the test: Modern language aptitude test. *The Personal Guidance Journal, 38*(7), 582–84.

Van Gelderen, A., Schoonen, R., de Glopper, K., Hulstijn, J., Snellings, P., Simis, A., & Stevenson, M. (2003). Roles of linguistic knowledge, metacognitive knowledge and processing speed in L3, L2 and L1 reading comprehension: A structural equation modeling approach. *International Journal of Bilingualism, 7*(1), 7–25. https://doi.org/10.1177/13670069030070010201

The Role of Language Aptitude in Second Language Attrition

Astrid Elisabeth Lehner

Abstract Only a few studies have aimed at identifying the role of aptitude in language attrition. Findings indicate that in early bilinguals, aptitude assumes a "compensatory function," allowing high-aptitude speakers to maintain and further develop a language in which they only receive very limited or no input, in contrast to lower-aptitude speakers. A correlation between L1 maintenance and language aptitude demonstrates the supportive role of lexical language learning aptitude for the retention of a language. Since prior research has found correlations between L1 and L2 proficiency and general language aptitude for early bilinguals, it is hypothesised that this association likewise holds true for late bilinguals. In the present study, no correlation is expected to be found between the frequency of L2 use and L2 proficiency scores within a group of speakers with an above-average language aptitude. However, it is expected that a longer time of non-usage of an L2 will lead to a lower L2 proficiency in below-average aptitude participants. Twenty-nine English native speakers who studied German as a foreign language were tested by means of a questionnaire, a language aptitude test, and proficiency tests for English and German. Results suggest that for these late bilinguals, only below-average aptitude learners depend on a longer duration of L2 learning, possibly also on earlier learning start dates, and on continuous input in the L2. Aptitude seems to enable above-average aptitude learners to achieve higher L2 levels within a shorter time and to prevent attrition. Finally, differences may also arise between males and females in L2 forgetting.

1 Introduction

Although second language (L2) or foreign language (FL) attrition shares many features of first language (L1) attrition, additional linguistic and extra-linguistic variables have to be taken into account when investigating the complex process of

A. E. Lehner (✉)
Department of English and American Studies, University of Vienna, Vienna, Austria
e-mail: astrid.lehner@aon.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_16

non-pathological forgetting of a second or foreign language (Bardovi-Harlig & Stringer, 2010; Schmid, 2011).[1] These variables, among others, include age of onset of L2 acquisition (AOA), age at which the acquisition process was stopped, level of peak L2 attainment, duration of contact with the L2, length of residence in a host country, motivation and personal attitudes towards language learning, a certain L2 and its speech community, literacy in the respective language, and a person's overall language learning aptitude (Bardovi-Harlig & Stringer, 2010). First of all, these factors play a crucial role in second language acquisition (for an overview, see Dörnyei, 2005). Some of these variables also constitute personality traits and/or cognitive factors which serve as a psychological basis for language learning and forgetting, and which account for both the great outcome variability in L2 learning (Dörnyei, 2005) and possibly L2 attrition (with not all linguistic subsystems affected to the same degree by attrition, see Wei, 2014). However, not all of these variables have received sufficient attention, with language learning aptitude being a prime example (Granena and Long 2012). So far, only a handful of studies have investigated language aptitude not only as a controlling, but as an indeed *key* variable in second language acquisition research (Granena & Long, 2012). Based on what has been investigated to date, general language learning aptitude has been shown to be positively correlated with native-like performance in L1 and L2 in early bilinguals (Bylund, Abrahamsson & Hyltenstam, 2012). In addition, a high language aptitude appears to be a requirement for native-like L2 proficiency in late bilinguals (Abrahamsson & Hyltenstam, 2008). Comparing aptitude to other variables, such as age of onset of acquisition or length of residence in a host country, aptitude apparently counterbalances effects of age and late onset of L2 acquisition, at least with regards to the lexical subsystem (Granena & Long, 2012; Abrahamsson & Hyltenstam, 2008).

Concerning language attrition, aptitude has hardly been investigated at all to date, with a study conducted by Bylund, Abrahamsson and Hyltenstam (2009) representing an important exception. These researchers have shown that in early bilinguals who do not use their L1 on a regular basis, speakers with an above-average language aptitude were more native-like in their grammatical intuition than speakers with an aptitude below average. In addition, the latter group's proficiency in the L1 was positively correlated with the amount of daily L1 use, a result which was not found to be the case in the above-average aptitude group. To that extent, Bylund, Abrahamsson and Hyltenstam (2009) concluded that language aptitude "has a compensatory function in situations of reduced L1 contact, in that the speaker's degree of aptitude to a certain extent regulates his/her dependency on L1 contact to achieve and maintain L1 proficiency" (p. 459). In other words, speakers with an above-average language aptitude do not depend on continued and regular input of a language in their development and retention of this language as much as below-average speakers do. Hence, aptitude can be said to prevent language attrition in early bilinguals.

[1] Both terms, L2 and FL, will henceforward be considered as synonymous.

Opitz (2011) investigated the relation between L1 attrition and L2 acquisition in late bilinguals who, as adults, had migrated to a country where their L2 is an official majority language. Opitz (2011) aimed at identifying a possible association between linguistic aptitude, L1 maintenance and L2 attainment, among other variables. To that end, the subjects of her study had to complete three out of five subtests of the English version of the Swansea Language Aptitude Test (LAT; Meara, Milton & Lorenzo-Dus, 2002; Meara, 2005). It is important to note here that the LAT did not test English, the bilinguals' L2, but artificial or rare language material unknown to the test-takers, while English was the language in which the material was presented (Opitz, 2011). The analysis revealed a significant correlation between the level of L1 maintenance and the word-learning subtest assessing lexical memory (LAT B). Even if Opitz (2011) is cautious about the findings regarding the aptitude test,[2] the LAT B result could still indicate a positive role of, at least, lexical language learning aptitude for the retention of an L1 in late bilinguals.

In a recent study, Bylund and Ramírez-Galan (2016) investigated the effect of aptitude on L1 maintenance in late bilinguals. As in Opitz' (2011) study, their subjects had migrated into a different language community as adults and had only started to acquire their L2 after arrival in the L2 host country. As a consequence, they had experienced reduced contact with their formerly fully developed L1. Subjects were tested on grammatical intuition and aptitude, and their scores were compared to those of an L1 predominantly monolingual control group. Results showed that the monolingual control group scored significantly higher on the grammatical intuition test than did the bilingual group. Furthermore, language aptitude did not play a significant role in the sense that it did not predict grammatical intuition. Furthermore, regarding correlation tests, only a non-significant tendency appeared between single aptitude subtest scores and L1 grammatical intuition. On that basis, Bylund and Ramírez-Galan (2016) conclude that aptitude does not influence L1 attrition in late bilinguals. Considering the very similarly designed study by Bylund, Abrahamsson and Hyltenstam (2009), which did confirm a compensatory function of aptitude in L1 attrition of early bilinguals, Bylund and Ramírez-Galan (2016) argue that the explanation for the very diverging findings from these two studies relates to differences in early and late bilinguals – more precisely, in their different respective ages at the onset of attrition. While early, pre-pubescent bilinguals such as those in Bylund, Abrahamsson and Hyltenstam's (2009) study had started to experience reduced L1 contact before the age where it is safe to assume that their L1 was fully developed, late bilinguals as in the more recent paper are believed to have completed their L1 acquisition process *before* experiencing attrition. Aptitude is thus discussed either as a variable that prevents attrition in early bilinguals, or as a variable that allows a more complete L1 acquisition of early bilinguals even in scenarios with reduced L1 contact (Bylund & Ramírez-Galan, 2016).

[2] Opitz (2011) reports significantly lower LAT B scores of the bilingual and L1 native control group, compared to the L2 native controls. Even if she argues that administration of the LAT B in English, which was the L2 of the bilingual group, is unlikely to have influenced the performance on this subtest, it is still conceivable that such a bias does exist.

Due to the absence of further studies which investigate the link between language aptitude and attrition, and especially given consideration of Bylund and Ramírez-Galan's (2016) results, it is not yet clear whether Bylund, Abrahamsson and Hyltenstam's (2009) findings can also be extended to late bilinguals with reduced contact with their L2 or FL. This paper tries to fill this gap with novel, empirical findings, which may then serve as a basis for further research.

The purpose of the present study is to investigate the role of language aptitude in late L2 or FL attrition. More precisely, the focus is on exploring whether differences exist in proficiency levels, language maintenance, and language attrition between high and low aptitude learners of German as a foreign language. For that purpose, the present study aims at replicating the findings of Bylund, Abrahamsson and Hyltenstam (2009) for late L2 acquisition. Two hypotheses have been formulated. The first hypothesis predicts that positive correlation exists between the degree of language aptitude and the L1 and L2 proficiency levels in late bilinguals, regardless of whether the L2 is being used frequently or not. Hence, it is assumed that there is no correlation between L2 proficiency and frequency of L2 use. Specifically, this means that a higher degree of aptitude results in a higher proficiency level, both of the L1 and the L2, while a lower degree of aptitude results in a lower proficiency level of L1 and L2.

The second hypothesis assumes a negative correlation between the length of the attrition period and L2 proficiency for below-average-aptitude speakers (length of attrition period is operationalized as time of reduced or no contact with the L2). At the same time, this hypothesis predicts that no such correlation exists between the length of the attrition period and L2 proficiency in above-average-aptitude speakers. In addition, hypothesis II predicts a negative correlation between aptitude and L2 attrition.

2 Methodology

2.1 Participants

In order to test the two hypotheses, participants with comparable linguistic and educational backgrounds were recruited. A first basic requirement was English as the first language. All participants were also required to have learnt German as a foreign language, preferably in a formal setting. Moreover, it can be said that most participants are late bilinguals in the sense that they had not learnt German before the onset of puberty (see Bylund, Abrahamsson & Hyltenstam, 2009).

The 29 participants, aged between 19 and 68 years (mean = 33.24), mainly come from the United States (17 participants), the United Kingdom (eight), the Republic of South Africa (four), and Australia (one). As for gender, 18 participants are female (62.1%) and 11 are male (37.9%).

Concerning educational background, the group is homogenous. That is, almost all participants hold a university degree, either at Bachelor's level (48.3%, 14 participants) or at Master's level (34.5%, ten), with one participant holding a doctoral degree and four participants having a high school diploma. Additionally, the highest qualification related to the German language was uniformly distributed. Most respondents completed university studies with German as a focus or the main language of instruction, such as Interpretation, Translation or German Studies (41.4% or 12 participants). In addition, the participants also hold an advanced German language certificate (CEFR level B1 and B2, "Independent user", 6.9% and 10.3% or two and three participants respectively; level C1 and C2, "Proficient user", 13.8% and 6.9% or four and two participants).[3]

As far as the language learning background is concerned, all participants have learnt several foreign languages, with a mean total of 3.97 foreign languages (minimum 1, maximum 20). Recalculation, with exclusion of one outlier ("20 foreign languages learned"), reveals a minimum from one to a maximum of nine foreign languages learnt (mean 3.39). The age of onset of L2 acquisition (AOA), i.e. German, ranges from 1 to 40 years ($n = 29$, mean = 15.52 years). Two participants who displayed a very early AOA, 1 and 3 years respectively, while the AOA for the remaining majority of participants ranges from 11 to 40 years ($n = 27$, mean 16.52). The possibility of the influence of AOA and some other factors on attrition was controlled in the analysis. Age of onset of L2 "non-learning," refers to the age at which the participants stopped learning German actively, ranges from 17 to 27 years ($n = 9$, mean = 22.56), with 20 participants claiming they were still learning German. To that effect, the length of L2 acquisition which equals the time between AOA and age of onset of L2 "non-learning" or current age in cases where participants are still learning is distributed between one and 57 years (mean = 14.25). The length of the period in which attrition can be assumed equals the length of time during which the participant claimed to have experienced reduced or no contact with German. This figure ranges from zero to 45 years, with a low mean of 3.38 years.

2.2 Instruments and Procedures

Participants had to complete an online questionnaire[4]and an array of computer-based tests in a pre-determined order of succession. Firstly, the questionnaire examined the personal, educational and linguistic backgrounds at hand. Secondly, L1 and L2 proficiency were determined. For this purpose, participants first had to take the English and then the German version of the LexTALE proficiency test (Lemhöfer & Broersma, 2012). Finally, aptitude was measured using parts III, IV and V of the Modern Language Aptitude Test (MLAT; Carroll & Sapon, 2002). Aptitude was

[3] For a detailed description of the Common European Framework for Languages (CEFR) see http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf

[4] Available on request

calculated as a compound score of the three sub-tests. Participants were provided with precise instructions and the corresponding links to the tests via email. Recruitment was achieved through personal contacts and with the help of the *Goethe Institut* in London.

3 Results

Statistical analysis was conducted using IBM SPSS Statistics version 24.0 (IBM Corporation, New York). In order to test the hypotheses for differences between participants with varying levels of language aptitude, group comparisons were required. That is, for various tests, the data was divided in accordance with the groups' median language aptitude score. Participants with an overall language aptitude score above the median constituted the *high-aptitude group*, while participants with an aptitude score below the median were subsumed in the *low-aptitude group*. However, since this division is based on this group-based relative measurement, the labels *high-* and *low-aptitude group* must not be understood as an absolute evaluation of the participants' language aptitude, but rather as a method which allows group comparison. For all calculations, a p-value of $p \leq .05$ was considered as threshold for statistical significance, while a value of $p \leq .001$ was considered to indicate high significance (Koller, 2014).

Attrition was calculated in two ways. Firstly, attrition was represented as the difference between self-reported German proficiency level at the time of peak and self-reported current German proficiency level, labelled *Difference German proficiency Self-Report Peak vs. Today*; secondly, attrition was calculated as a compound score. For tests which included the compound attrition score, six participants had to be excluded due to missing values (thus, $n = 23$ for these tests).

Variables were tested for normality of distribution with the Kolmogorov-Smirnov and Shapiro-Wilk-Test. The variables *AOA of German, Overall language aptitude score*, *German language proficiency score*, *Compound attrition score,* and *MLAT IV Score* were normally distributed. Other variables such as *English language proficiency score* or *Exposure to German* were not normally distributed and were analysed using non-parametric tests accordingly.

3.1 Hypothesis I – Aptitude and L1 Proficiency

Regarding the connection between aptitude and language proficiency, the Spearman-Rho correlation analysis shows a possible tendency towards a positive correlation of $r_S = .341$ between the L1, i.e. English, proficiency score, and the overall language aptitude score in the low-aptitude group. However, this result is not significant (p = .336). For the high-aptitude group, the analysis reveals no correlation between L1, i.e. English, proficiency scores and a participant's overall language aptitude score (r_S = .092, p = .765).

3.2 Hypothesis I – Aptitude and L2 Proficiency

The correlation analysis again reveals a tendency towards a positive correlation between L2, i.e. German, proficiency and overall language aptitude scores within the group with an aptitude above the median (r = .481, p = .096). Regarding the group with an aptitude below the median, the correlation coefficient of r = .158 shows practically no correlation between the German proficiency and language aptitude scores. This result is not significant (p = .663).

3.3 Hypothesis I – Group Differences in L1 and L2 Proficiency

Furthermore, hypothesis I predicts differences in the L1 and L2 proficiency scores between the low and high-aptitude group. For the variable *German language proficiency score*, the Levene's test was not significant (p = .864), so variances are assumed to be homogeneous. Subsequently, the t-test demonstrates that there is a significant difference in the mean German proficiency scores between the low- and high-aptitude group (p = .046). For the variable *English language proficiency score*, the Levene's test was also not significant (p = .872), as was the t-test (p = .671).

In addition to the t-test, Fig. 1 below visually presents the L1 and L2 proficiency scores for both groups. It indicates a higher range of L2 proficiency scores in the high-aptitude group (mean .764, SD .104, minimum score 58.75%, maximum score 100%). In the case of the low-aptitude group, the mean L2 score is .675, with a standard deviation of .095, a minimum L2 proficiency score of 55% and a maximum of 86.25%. Meanwhile, for L1 scores, the high-aptitude group's mean is .956 (SD .102, minimum score 62.5%, maximum 100%), while the mean for the low-aptitude group is .940 (SD .061, minimum score 85%, maximum 100%). Thus, interestingly, the lowest L1 proficiency score from the low-aptitude group is higher than the respective score from the high-aptitude group. However, the t-test and Mann-Whitney-U-test do not yield any significant differences between the groups in relation to L1 proficiency test scores.

The 95% confidence interval error bars in the diagram overlap in the case of both the L1 and L2 scores. This might lead to the assumption that, despite the results of the t-test, a significant difference in L2 scores between both groups cannot arise. However, it has to be kept in mind that an overlap of 95% CI error bars does not necessarily equal non-significance. This is different from an overlap of standard error bars, which does equal non-significance (see Motulsky, 2002).

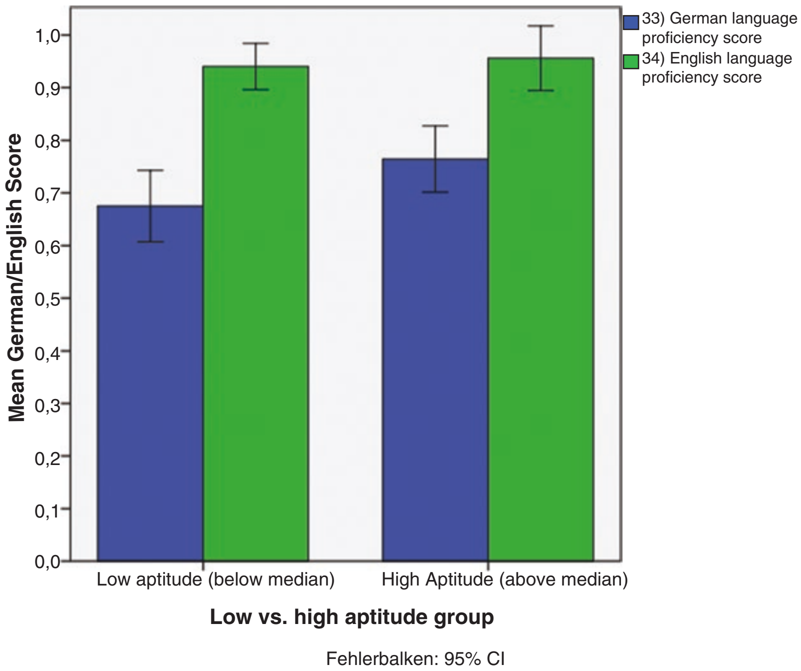


Fig. 1 Group comparison of L1 and L2 scores

Since the sample size of n = 23 was rather small for a t-test, and given the overlap of the 95% CI error bars in the diagram, a Mann-Whitney-U-test was also performed in order to double-check the findings (Koller, 2014). The Mann-Whitney-U-test confirms the findings from the t-test, and shows that the overlap in the 95% CI error bars is not relevant. Specifically, the Mann-Whitney-U-test reveals firstly that the difference in the mean L2 proficiency scores between the two groups is significant ($p = .037$), and secondly that the difference in the mean L1 scores is not significant between both groups ($p = .182$).

3.4 *Hypothesis I – L1 and L2 Proficiency, Frequency of L2 Use and Age of Onset of L2 Learning*

A third aspect of hypothesis I concerns the relationship between L2 proficiency scores and the frequency of L2 use. The Spearman correlation analysis of the sample reveals that for both groups, no significant correlation appears between the German proficiency score and self-reported exposure to German ($r_S = .252$, $p = .547$).

Regarding the age of onset of L2 acquisition and its possible influence on L2 proficiency, the analysis reveals that for the high-aptitude group, no correlation exists between the variables ($r = -.049$, $p = .874$). For the low-aptitude group, a tendency towards a negative correlation between the age of onset of L2 learning and the German proficiency score ($r = -.392$, $p = .263$) could emerge. However, this is not significant.

Concerning the duration of German learning, no correlation with German proficiency for the high-aptitude group ($r_S = .092$, $p = .765$) is seen. In contrast, for the low-aptitude group, a significant, very strong, positive correlation ($r_S = .740$, $p = .014$) surfaces.

3.5 Hypothesis II: Aptitude and L2 Attrition

The second hypothesis primarily examines the relationship between general language aptitude and language attrition. Since it was not possible to conduct a longitudinal study with multiple measurements of L2 German proficiency, the analysis has to rely on self-reported language levels at the time of personal peak attainment in the L2.[5] In order to test the reliability of self-reported answers regarding former and current German proficiency, a correlation test was conducted between self-reported current German proficiency level and actual performance on the proficiency test. The analysis shows a positive, significant correlation of the two variables ($r_S = .438$, $p = .017$), thus indicating validity of the measurement and reliability of participants' responses to this item.

Attrition was, however, not only calculated from the difference between self-reported peak and current L2 proficiency score, but likewise as a compound score derived from answers to questionnaire items which aimed at revealing whether a participant considered his or her German skills to be undergoing attrition, and/or the self-perceived extent to which German was forgotten. This variable thus carries the notion of a degree of attrition. In order to check internal reliability, Cronbach's Alpha was calculated for these items. The score of $r_\alpha = .714$ reveals a sufficient degree of reliability, since it is above .7 (Kline, 1999, quoted by Wucherer, 2015). In addition, a Spearman correlation test between the compound and self-reported attrition scores was run to further check their reliability. The result is a highly significant and very strong, positive correlation between both variables ($r_S = .800$, $p < .001$). The measurement tools employed thus seem to be valid, and the two attrition scores and respondents' answers appear to be reliable.

Turning to the actual hypothesis testing as for aptitude and its connection to attrition, a Spearman Rho correlation test applied to the whole sample reveals a non-significant, moderate, negative correlation between self-reported attrition and the overall language aptitude score ($r_S = -0.222$, $p = .308$).

[5] A self-reported attrition score was calculated as the difference between self-reported peak and current German proficiency levels. Both items were asked in the questionnaire and to be answered on a Likert scale from 0 to 10.

A separate analysis of both groups shows that in the high-aptitude group, a tendency most likely occurs towards a moderate, negative correlation between both variables; however, this result is not reliable ($r_S = -0.365$, $p = .220$). Concerning the low-aptitude group, no correlation between the variables arises, which is a result with a high error probability ($r_S = 0.082$, $p = .821$).

In order to double-check these findings, Pearson correlation tests were conducted with the compound attrition score. Regarding the whole sample, no significant correlation between the compound attrition and the overall language aptitude scores emerged. However, a slight tendency towards a negative correlation ($r = -.157$, $p = .476$) may occur.

Moreover, scrutiny of the results for both groups reveals no significant correlation between the two variables. It is however noteworthy that for the high-aptitude group, the algebraic sign is negative, therefore indicating a possible negative correlation ($r = -.440$, $p = .132$) between attrition and aptitude, while the opposite is true for the low-aptitude group ($r = .355$, $p = .315$).

Finally, it is clear that both variables, the self-reported and compound attrition score, yield comparable, non-significant, results.

3.6 Hypothesis II: Length of Attrition Period, Aptitude, and Attrition

Regarding connections between aptitude, self-reported L2 attrition and the length of reduced contact with the L2 (which is the length of time of assumed attrition) for the high-aptitude group, the Spearman test revealed a possible tendency towards a positive correlation between self-reported attrition and length of attrition time ($r_S = .367$, $p = .218$). In other words, this could indicate that the longer the time of reduced contact with the L2, the higher the self-reported difference between proficiency levels at the time of personal peak and time of testing. This effect is also evident for the low-aptitude group. That is, even if insignificant, a positive correlation between the variables ($r_S = .365$, $p = .314$) may nevertheless prove evident.

When using the compound attrition score to test the same effect, the results are partly significant. For the high-aptitude group, the correlation coefficient is $r_S = .331$ with an error probability of $p = .270$, which again might be a tendency. However, for the low-aptitude group, a significant result ($r_S = .755$, $p = .012$) stands out.

The same effect was also sought in a correlation test between the length of time with reduced or no contact with German and the German proficiency score. As predicted by hypothesis II, no significant correlation appeared between these variables within the high-aptitude group ($r_S = -.184$, $p = .548$). Contrary to the expectations for the low-aptitude group, no significant correlation arose in this case either ($r = .259$, $p = .469$).

Furthermore, the correlations between overall language aptitude and the self-reported extent to which German has been forgotten, or begun to be forgotten, were also tested. The analysis shows that for the high-aptitude group, a non-significant, negative correlation between these variables ($r_S = -.402$, $p = .173$) emerges. Regarding the low-aptitude group, the same, non-significant connection is found, but instead of a negative correlation, it is instead positive ($r_S = .278$, $p = .437$).

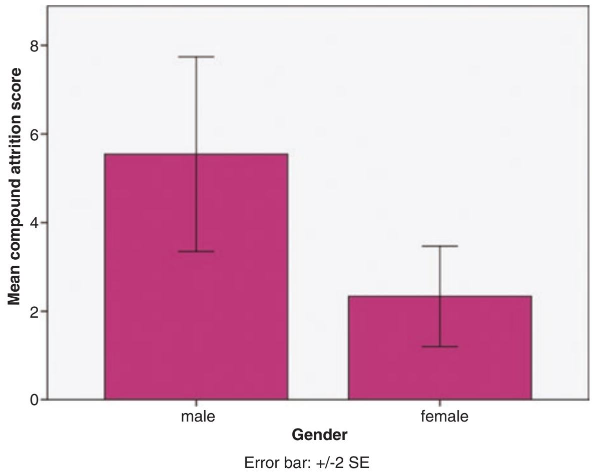
Furthermore, the extent to which L2 German was self-reportedly forgotten significantly correlates with the length of time with reduced or no contact with this language. For the high-aptitude group, the Spearman correlation coefficient is not significant ($r_S = .235$, $p = .439$), while for the low-aptitude group, the correlation is highly significant and strong.

($r_S = .826$ $p = .003$).

3.7 Additional Tests: Influence of Age, Gender, AOA German and Highest Educational Degree

Additional tests were conducted in order to control the variables *Age*, G*ender, AOA German* and *Highest educational degree* and their possible connection to, or influence on, attrition and aptitude. For these calculations, the metric and normally distributed *Compound attrition score* was used. Non-parametric Chi-Square-tests had to be applied for checking the variables *Age*, *Gender* and *Highest educational degree*. For the metric *AOA German*, a Pearson correlation test was run. None of these tests were significant, thus yielding that neither age, nor gender, nor AOA German, nor the educational degree seem to have a relevant correlation with attrition. What's more, aptitude was not correlated with these variables, as had to be expected. However, the p-value for gender ($p = .071$) may well indicate a tendency. A Mann-Whitney-U-test confirmed significant differences in the compound attrition score between males and females ($p = .014$). The mean compound attrition score is 5.55 for males and 2.33 for females. Yet, as evident in Fig. 2, standard error

Fig. 2 Gender and attrition



bars slightly overlap, which again would indicate non-significance. It might possibly be true that women and men remember or forget foreign languages differently. However, since the role of gender for attrition was not a key question in this project, it will not be treated more extensively here and must be left open for future research.

4 Discussion

Since no significant correlation arises between the overall aptitude score and English proficiency scores of the high- and the low-aptitude group, this part of hypothesis I cannot be confirmed. There might be a tendency towards a moderate, positive correlation for the low-aptitude group, however, this is not significant ($r_S = .341$, $p = .336$). This means that especially in the case of lower-aptitude learners, the benefits attached to learning foreign languages should manifest twofold: they would not only be learning the new language, but also further developing their first language at the same time. However, more testing is needed to confirm this tentative interpretation.

Regarding proficiency in late additional languages, the aptitude of speakers with a higher aptitude almost significantly contributed to higher proficiency scores ($r = .481$, $p = .096$). This finding could confirm the respective part of hypothesis I, while also corroborating Bylund and colleagues' findings (2009) and findings by Abrahamsson and Hyltenstam (2008). A higher language aptitude seems to enable language learners to acquire a foreign language to a higher proficiency level, compared to learners with a lower language aptitude. However, further testing will be needed for late L2 learners. Comparison of L2 German proficiency scores of the high-aptitude group with proficiency scores of a German native control group is advisable for future research. A different aptitude test would have to be employed for this purpose, since the MLAT depends on a very good command of English. Hence, when used for non-native-speakers of English, the test results can be biased due to varying English proficiency levels.

4.1 Hypothesis I – Group Differences in L1 and L2 Proficiency

The result of the t-test shows that no significant group difference arises in the mean L1 proficiency score. Nevertheless, a significant difference in the mean L2 proficiency scores emerges between the low- and high-aptitude groups ($p = .046$). This finding corroborates results from earlier studies which have shown a positive effect of language aptitude on L2 learning (DeKeyser, 2000, quoted by Bylund, Abrahamsson & Hyltenstam, 2009). It seems that aptitude is more visible in L2 proficiency than in L1 proficiency, at least in late bilinguals. A reasonable conclusion therefore holds that the assumption of hypothesis I - which predicts group differences in the mean proficiency scores - can be validated for L2 but not for L1 proficiency.

4.2 Hypothesis I – L1 and L2 Proficiency, Frequency of L2 Use and Age of Onset of L2 Learning

The relation between L1/2 proficiency and frequency of L2 use was also investigated by Bylund, Abrahamsson and Hyltenstam (2009) in their study on prepubescent L1 attrition. They found that for participants with an aptitude below the mean, self-reported daily L1 use significantly, positively correlated with grammaticality judgement scores. In other words, "daily L1 use had a positive effect on GJT performance" (Bylund et al., 2009, p. 455). Yet, no such correlation was found for participants with an aptitude above average (455 f.). Since the present study did not find any correlation between frequency of L2 use and L2 proficiency, the findings of Bylund et al. (2009) concerning speakers with an aptitude below average cannot be replicated. Yet, for the high-aptitude group, the current study nevertheless found the same non-correlation between the amount of exposure to the language presumably having undergone attrition, and scores indicating proficiency in that language. This substantiates the findings concerning speakers with a higher language aptitude from earlier research (Bylund et al., 2009).

Regarding the influence of age of onset of German learning on German proficiency scores, no correlation was found for the high-aptitude group ($r = -.049$, $p = .874$), while for the low-aptitude group, a tendency towards a negative correlation ($r = -.392$, $p = .263$) surfaces. This points at a possibly higher dependency of learners with a lower language aptitude on earlier learning start dates, compared to learners with a higher aptitude. Henceforth, age of onset of L2 learning seems to play a more decisive role for learners with a lower degree of language aptitude. On the contrary, for learners with a higher aptitude, aptitude appears to counterbalance negative effects of late age of onset, a finding which corroborates earlier results (Granena & Long, 2012; Abrahamsson & Hyltenstam, 2008). Since the finding is insignificant, however, further research is needed.

This point is further substantiated by a significant, positive correlation for the low-aptitude group between the duration of German learning and the German proficiency score ($r_S = .740$, $p = .014$). No such correlation was found for the higher-aptitude group. This firstly suggests that for persons with a higher language aptitude, neither age of onset of L2 learning, nor the total duration of L2 learning, significantly contribute to the ultimate L2 proficiency. Secondly, the findings could suggest that low-aptitude learners require a longer duration of language acquisition. This entails a potential advantage for this group of learners if their language acquisition starts relatively early.[6]

In short, the predicted positive correlation between German proficiency scores and the total amount of exposure to German in the low-aptitude group was found, while, as expected, no such correlation was found for the high-aptitude group.[7]

[6] Obviously, "earlier" onset of language acquisition in this context of late bilingualism still refers to a relatively late AOA, compared to (bilingual) first language acquisition.

[7] "Total amount of exposure to German" here means exposure in the sense of learning duration. It does not equal the item "Exposure to German" from the questionnaire, which refers to current exposure to the language.

4.3 Hypothesis II: Aptitude, L2 Attrition, and Length of Attrition Time

A core question of the present research project concerns the relationship between language aptitude and second language attrition. Hypothesis II predicts a positive correlation between these variables, however, no significant correlation was found either for the whole sample, or for the two groups separately. Only a possible, slight tendency towards a negative correlation between the compound attrition score/self-reported attrition score and overall language aptitude - which was nevertheless insignificant - could be found for the high-aptitude group and the whole sample respectively. Even if non-significant ($p = .220$ and $p = .132$) for the high-aptitude group, these findings could nevertheless hint at a negative correlation between aptitude and attrition ($r_S = -.365$, $r = -.440$). The same applies to the whole sample.

Another prediction of hypothesis II relates to the influence of the length of time in which reduced or no contact with the L2 was experienced by participants. However, no significant result was found for the self-reported attrition score for both groups. Calculating with the compound attrition score, no significant result arose for the high-aptitude group, but instead for the low-aptitude group ($r_S = .755$, $p = .012$). This means that for the lower-aptitude group, the length of time with reduced or no contact with the L2 has a stronger impact on language attrition than for the high-aptitude group. In other words, aptitude seems to outweigh at least to a certain degree the negative effect of a longer time without or with only limited L2 contact. This again substantiates findings by Granena and Long (2012).

When looking at German proficiency and length of attrition time, no correlations for the high-aptitude group emerged, as was predicted by hypothesis II. Nonetheless, the assumption of such correlation for the low-aptitude group could, unexpectedly, not be upheld by the results ($r_S = .259$, $p = .469$). Thus, only a part of hypothesis II can be validated.

Regarding the extent to which German was self-reportedly forgotten, and its relation to language aptitude, no statistically significant result was found. Yet, there might again be a tendency of high-aptitude speakers towards a negative correlation of their aptitude score and the extent to which their L2 was forgotten ($r_S = -.402$, $p = .173$). For the low-aptitude group, no correlation was found; it is however interesting that the algebraic sign was positive for this group, compared to the high-aptitude group ($r_S = .278$, $p = .437$). Thus, although insignificant, this finding is interesting since the difference in the algebraic sign could indicate a qualitative difference between the two groups regarding aptitude and the degree of language attrition. As might be expected, the tendency within the high-aptitude group yielded the following - the higher the aptitude, the smaller the extent to which German was forgotten. This could point to the role of language aptitude for the maintenance of a language, which was shown to be a decisive one by Bylund and his colleagues (2009). They concluded that aptitude seems to compensate for reduced input in a language, thus helping a speaker to not forget that language (Bylund, Abrahamsson & Hyltenstam, 2009).

Further substantiating this point is the fact that the extent to which German was forgotten significantly correlated with the length of attrition time, at least for the low-aptitude group ($r_S = .826$, $p = .003$; high-aptitude group $r_S = .235$, $p = .439$). Apparently, language learners with a lower aptitude seem to forget a language to a greater extent than the learners or speakers with a higher aptitude, and it seems that this extent grows with a larger time period in which reduced or no contact with the language was experienced. Clearly, this finding corroborates the results from Bylund, Abrahamsson and Hyltenstam (2009), as lower-aptitude speakers seem to depend more on continuous language input in order to further develop or maintain competence in this language. Higher language aptitude seems to indeed prevent attrition and maintenance of an L2.

Similar to Bylund and his colleagues' findings, which demonstrate a correlation for participants with a lower-than-average language aptitude and not for those with an aptitude above average, the results from the present study also suggest that only lower-aptitude learners or speakers of an L2 depend on continuous input in that language in order to prevent forgetting it. Higher-aptitude speakers benefit from their aptitude, which seems to lessen the effect of non-frequent use of, or contact with, the L2 (see Bylund, Abrahamsson & Hyltenstam, 2009).

Since the sample size in the present study is rather small, further testing could reveal whether the above observation likewise holds true for high-aptitude learners or speakers as the data from the present study suggest. Yet, if a significant correlation between the length of the attrition period and the extent to which the L2 was forgotten was also found for the higher-aptitude group, this would indicate that not only the lower-aptitude but also the high-aptitude speakers with a late age of onset of L2 learning depend on ongoing input in the given language. In this case, two reasons could be at play here. Firstly, since the differentiation between the high and low-aptitude participants in this study was done on the basis of the median of the aptitude score of the whole sample, it might, in terms of raw aptitude scores, qualitatively differ from the high-aptitude participants from Bylund, Abrahamsson and Hyltenstam's study (2009). A second explanation could be that there is indeed a difference for early and late bilinguals: Bylund and his colleagues (2009) tested early bilinguals, while the present study concerns late bilinguals who mostly learnt German after the age of 10 years. It could therefore indeed be the case that late bilinguals, even if they have a higher-than-average language aptitude, depend on continuous input in their late language(s). In contrast, early bilinguals seem to depend to a lesser extent on such input if they have a sufficiently high aptitude (see Bylund, Abrahamsson & Hyltenstam, 2009). As mentioned above, the more recent study by Bylund and Ramírez-Galan (2016) indeed found a qualitative difference in late bilinguals' L1 attrition, compared to early bilinguals'. Future research could aim at directly comparing late and early bilinguals with high language aptitude scores to clarify this issue. As was shown, the present data suggest that the length of attrition time does not influence the extent of L2 forgetting in higher aptitude speakers, but indeed in lower aptitude speakers.

5 Conclusion

The present project investigated the relationship between language aptitude and second language acquisition and maintenance. While some findings have substantiated results from earlier research, other findings were rather unexpected. For instance, the apparent lack of an association between language aptitude and L1 proficiency scores was, based on earlier research, not predicted. It seems plausible that a larger sample size would yield results that can corroborate Bylund, Abrahamsson and Hyltenstam's findings in this regard (2009).

However, significant differences in L2 proficiency scores between participants with a lower and a higher general language aptitude, and a higher mean L2 score for participants with an aptitude above the median, substantiate earlier findings of aptitude playing a decisive role in acquiring a high proficiency level in an L2 (Abrahamsson & Hyltenstam, 2008).

In addition, only the lower-aptitude group of late bilinguals seems to depend more on earlier learning start dates, a longer duration of L2 learning and on continuous input in the language in order not to forget it. This conclusion is drawn based on significant correlations between the following variables: L2 proficiency test score and duration of L2 learning, length of attrition time and compound attrition score, length of attrition time and extent of L2 attrition, and tendency towards a negative correlation of age of onset of L2 learning and the L2 proficiency test score. On the other hand, since such correlations have not been found for the higher-aptitude group, a higher general language aptitude seems to not only enable foreign language learners to achieve a higher proficiency level in the L2 within a shorter time, but to likewise maintain the L2 even in times of reduced or infrequent contact with this language. In sum, these results suggest that earlier findings for early bilingualism (Bylund, Abrahamsson & Hyltenstam, 2009) can also be assumed for late bilingualism. The question of whether gender also influences the forgetting of an L2 however remains open for future research.

References

Abrahamsson, N., & Hyltenstam, K. (2008). The Robustness Of Aptitude Effects In Near-Native Second Language Acquisition. *Studies in Second Language Acquisition, 30*(4), 481–509. https://doi.org/10.1017/S027226310808073X

Bardovi-Harlig, K., & Stringer, D. (2010). Variables in second language attrition: Advancing the state of the art. *Studies in Second Language Acquisition, 32*, 1–45. https://doi.org/10.1017/S0272263109990246

Bylund, E., & Ramírez-Galan, P. (2016). Language aptitude in first language attrition: A study on late Spanish-Swedish bilinguals. *Applied Linguistics, 37*(5), 621–638. https://doi.org/10.1093/applin/amu055

Bylund, E., Abrahmasson, N., & Hyltenstam, K. (2012). Does first language maintenance hamper nativelikeness in a second language? A study of ultimate attainment in early bilinguals. *Studies in Second Language Acquisition, 34*(2), 215–241. https://doi.org/10.1017/S0272263112000034

Bylund, E., & Ramírez-Galan, P. (2016). Language aptitude in first language attrition: A study on late Spanish-Swedish bilinguals. *Applied Linguistics, 37*(5), 621–638. https://doi.org/10.1093/applin/amu055

Carroll, J., & Sapon, S. (2002). *Modern language aptitude test*. Bethesda, MD: Second Language Testing.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*(4), 499–533.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. New York: Routledge.

Granena, G., & Long, M. H. (2012). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*(3), 311–343. https://doi.org/10.1177/0267658312461497

Kline, P. (1999). *The handbook of psychological testing*. London: Routledge.

Koller, M. M. (2014). *Statistik für Pflege- und Gesundheitsberufe*. Vienna, Austria: Facultas.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*, 325–343. https://doi.org/10.3758/s13428-011-0146-0

Meara, P. M. (2005). *Llama language aptitude tests: The manual*. Swansea: University of Wales Retrieved from http://www.lognostics.co.uk/tools/llama/llama_manual.pdf. Accessed 19 July 2017

Meara, P. M., Milton, J. L., & Lorenzo-Dus, N. (2002). *Language aptitude tests*. Newbury, UK: Express.

Motulsky, H. (2002). *The link between error bars and statistical significance*. GraphPad Software Retrieved from https://egret.psychol.cam.ac.uk/statistics/local_copies_of_sources_Cardinal_and_Aitken_ANOVA/errorbars.htm. Accessed 19 Sept 2016.

Opitz, C. (2011). *First language attrition and second language acquisition in a second language environment*. Dublin, Ireland: Trinity College Centre for Language and Communication Studies Retrieved from http://www.let.rug.nl/languageattrition/Papers/Opitzthesis. Accessed 7 Nov 2015

Schmid, M. S. (2011). *Language attrition*. Cambridge, UK: Cambridge University Press.

Wei, J. (2014). Selectivity of second language attrition. *Theory and practice in language studies, 4*(8), 1603–1608. https://doi.org/10.4304/tpls.4.8.1603-1608

Wucherer, B. V. (2015). Language is a girlie thing, is it not? Sex/gender differences in the limelight: An interdisciplinary approach. Diploma thesis, Department of English. Vienna: University of Vienna.

So What's the Deal Now!? Am I Talented or Not?

Stefanie Rüdegger

Abstract In research on language aptitude various ways of measuring individual differences in language abilities have been developed. Yet, it could be presumed that two commonly used aptitude tests will show similar results. However, in this investigation it cannot be supported that the aptitude tests under consideration reveal similar results. Language aptitude, therefore, is a multi-componential phenomenon consisting of multiple independent cognitive abilities which apparently are not fully covered by current language aptitude tests in use. The study thus reveals that it may be important to firstly define an *aptitude of interest* before applying aptitude tests in order to receive valid aptitude information. Furthermore, it aims at raising awareness of the practical use of aptitude tests and intends to warn of drawing too general conclusions from aptitude test results.

1 Introduction

Language learning has always been one of the central issues of mankind which has led to multiple controversial theories which all seek to entangle the mystery of the acquisition of our native language. During the last decades the scientific interest, however, seems to have focused on second language learning in much detail. Apparently, the ability to master foreign languages has become of indispensable advantage to meet the demands of today's globalized world. Different linguistic theories all seek to explain the process of second language learning. What most of them though agree upon is emphasizing the differences of the individual language learners. Therefore, variables such as motivation, learning style, personality, attitudes, and personal beliefs of the language learner are all

S. Rüdegger (✉)
Department of Linguistics, University of Vienna, Vienna, Austria

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_17

related to second language learning outcomes and aim at providing justification to a predictable success in learning a foreign language (Lightbrown & Spada, 2013, p. 77).

Another important factor that has been investigated to find differences in foreign learning achievement is that of language learning aptitude. In fact, some people learn a foreign language faster, more effortless, and with apparently better results than others. To measure such talent and therefore to predict success in language learning, various language aptitude tests have been developed throughout the last 60 years.

This investigation deals with several language learning aptitude tests and their comparisons. Moreover, it presents the findings of a cross-sectional study that was conducted to reveal if such a complex, multi-componential, and non-unitary ability as language learning aptitude can, in fact, be reliably measured by two aptitude tests under consideration.

1.1 Defining Foreign Language Aptitude

Apparently, some seem to learn a foreign language more easily than others. Foreign language aptitude thus seems to be a special, inherent talent available to people prior to learning, and best predicting success in foreign language (Ellis, 2008, p. 472). language aptitude is not about the ability to learn a foreign language but rather indicates the rate of progress of the learner under optimal learning conditions (Dörnyei & Ryan, 2015, p. 38). However, not even experts in language, language teaching, and mainstream psychology would dare to ultimately define what exactly language aptitude is. It thus appears rather impossible to provide one single definition of aptitude since "such a definition depends largely on both the theoretical and empirical context of a given author" (Nardo & Reiterer, 2009, p. 213). Most scholars however agree on several features related to aptitude (based on Jorgenson, 2008; Rysiewicz, 2003, p. 572):

- it is regarded as something special, or rather an *exceptional capability* in a given domain;
- it is regarded as a *potential*, e.g. something capable of development.
- it is an autonomous dimension independent of both, affective (anxiety, motivation, attitudes) as well as general cognitive factors;
- it is independent of academic ability or intelligence, although it partially overlaps with these domains;
- it is relatively stable over longer periods of time; not dependent on prior learning experience; not easily modifiable through training;
- it is not a single, unitary capacity but a composite of several relatively independent cognitive abilities (componential/multi-factor structure);
- it is always a better prognostic of L2 learning success than any other ID taken singly or in combination with each other.

1.2 Language Aptitude Tests

The concept of language aptitude was originally developed in the 1920s as an attempt to lower costs of language education in the USA (Thompson, 2013, p. 686). Back then aptitude tests were mainly used by the US Army during World War II as the military quickly needed people who most likely perform well in languages to keep language programmes cost-effectively (Sasaki, 2012, p. 317).

Coming from this more or less economic-political background, research in the field of language aptitude has found its grounds, and within time, shifted to a psycho-educational interest. Researchers started to develop aptitude tests to support psychologists and educators in their work since "tests can help psychologists and educators (a) know to whom to devote what levels and what kinds of resources, (b) be able to predict success in language learning instruction, and (c) be able to compare actual achievement with the achievement one might expect on the basis of foreign language learning ability" (Grigorenko, Sternberg, & Ehrman, 2000, p. 391).

Language aptitude testing and thus generating aptitude profiles helps teachers to adapt their teaching to the needs of their students, and therefore to accommodate and support them in their process of learning in the most efficient way.

Based on different conceptualizations different aptitude tests have been developed in the past decades. Four of them are briefly outlined and compared in the following section.

1.2.1 The Modern Language Aptitude Test

The most popular and still widely used test on language aptitude is the Modern English Aptitude Test (MLAT). It was developed by John B. Carroll together with his colleague Stanley M. Sapon, two of the most important researchers in language aptitude research. The MLAT can be seen as the established standard for aptitude measurement so that "newer tests seem to only complement the MLAT rather than supersede it, and their validation tends to include the MLAT as 'the benchmark'" (Grigorenko et al., 2000, p. 397).

The test was firstly published in 1959 and consists of five subtests. Each of the subtests intends to measure a different skill related to second language learning (Carroll & Sapon, 1958). Based on their research, Carroll and Sapon came up with four major components of foreign language aptitude:

(a) *phonetic coding ability*: the ability to identify distinct sounds, to form associations between those sounds and symbols representing them and to retain these associations,
(b) *grammatical sensitivity*: the ability to recognize the grammatical functions of words (or other linguistic entities) in sentence structures;

(c) *rote learning ability for foreign language materials*: the ability to learn associations between sounds and meanings rapidly and efficiently and to retain these associations; and
(d) *inductive language learning ability*: the ability to infer or induce the rules governing a set of language materials given samples of language materials that permit such inferences" (Sasaki, 2012, p. 316).

Based on this, aptitude construct the MLAT was developed. It was designed for adults who are literate with English as their first language. Strikingly, the test has not changed since it was first published and is still widely used today for various kinds of research purposes (Sasaki, 2012, p. 315).

The MLAT has fairly predictive power *(r = .4 to .6)* and therefore "predicts L2 learning success relatively well for both formal and informal L2 learning across different skills […]" (Sasaki, 2012) Three possible reasons for "this remarkable longevity" he continues are: "(a) the authors foresight, (b) the fact that the MLATs development was guided by ample empirical data collected from different types of educational settings using different types of teaching methods, and (c) because of the relatively slow development of language aptitude research"(p. 317).

1.2.2 The CANAL-F

Another language aptitude test- grounded in cognitive theory and very different to the MLAT- was developed by Elena L. Grigorenko, Robert J. Sternberg, and Madeleine E. Ehrman around the year 2000. It is called CANAL-F, which stands for Cognitive Ability for Novelty in Acquisition of Language- Foreign and is based on the idea that "one of the central abilities required for foreign language acquisition is the ability to cope with novelty and ambiguity" (Grigorenko et al., 2000, p. 392). The CANAL-F was developed to consider five acquisition processes running when learning a foreign language. Those are: selective encoding, accidental encoding, selective comparison, selective transfer, and selective combination. As the FL learner has to consistently deal with new linguistic material, "he or she has to decide where to focus his or her attention and how to use these processes accordingly" (Grigorenko et al., 2000, p. 392). These "processes operate at four levels of (a) lexis: dealing with one's learning, understanding, and use of words; (b) morphology: dealing with the words' structures and derivations; (c) semantics: dealing with one's understanding and use of the meaning of the words, based on information from the higher order units into which the words combine, such as sentences and paragraphs; and (d) syntax: dealing with one's learning, understanding, and use of the grammatical principles of organization that connect the words to the higher order units" (Grigorenko et al., 2000, p. 393).

The CANAL-F expects the participant to learn elements of an invented language. This new language, named *Ursulu*, is presented gradually, meaning in the beginning of the test participants do not know it at all, whereas in the end they will know enough *Ursulu* to understand a little story. It consists of five sections which deal with: "learning meanings of neologisms from context; understanding the meaning of passages; continuous paired-associate learning; sentential inference; and learning language rules" (Grigorenko et al., 2000, p. 394ff).

1.2.3 The LLAMA

The LLAMA aptitude test has been developed by Paul Meara (2005) and is an ongoing project of him and his students of English Language and Linguistics at the University of Wales Swansea. It is largely based on the MLAT but adds two aptitude dimensions which are not present in the MLAT. The LLAMA is a computer based test and is- in contrast to the MLAT- designed without the first language being a requirement (Meara, 2005).

1.2.4 The Hi-LAB

Another just recently developed and so far unpublished language learning aptitude test is the Hi-LAB, which stands for High Level Language Aptitude Battery. The test has been developed by the *Center for Advanced Study of Languages* (CASL) of the University of Maryland, and "unlike currently used aptitude tests, which predict success in the early stages of language learning, CASL's Hi-LAB uses innovative behavioural tasks to predict ultimate attainment" (Doughty, 2013).

The developers claim that the test has been developed because "the U.S. government has an urgent and growing need for foreign language professionals with high-level proficiency, but few learners who begin after age 12 are able to reach these levels" (Doughty, 2013). It is "a composite set of tests that measures cognitive and perceptual abilities designed to predict aptitude for learning foreign language to advanced levels. Using innovative measures to identify high-level language learners, CASL's Hi-LAB assists with the selection, hiring, and training of top language professionals" (Doughty, 2013). She proceeds that "the individual sub-tests of CASL's Hi-LAB battery have been combined into various composite scores, in order to provide information on a range of dimensions of language learning. All language learning involves processing rich and varied *input* from the target language, *interaction* of the learner with other speakers and with a variety of tasks in the language, and the processing of *feedback* in order to refine language proficiency to be more target-like" (Doughty, 2014). Factors, which are considered to have an impact on input-processing are: variability, authenticity, pattern learning, meaning association, and phonological perception. In terms of interaction, the Hi-LAB investigates the levels of attention switching and scaffolding. The third factor involved in language learning is the processing of feedback. The test intends to find out which type of feedback the particular language learner will find most effective. Unfortunately, the Hi-LAB is not accessible and further analysis cannot be presented here. However, the Hi-LAB looks fairly different from the first published aptitude test, the MLAT. Dörnyei (2015) claims that the Hi-LAB "represents a richer, theoretically grounded conceptualization of language learning ability" (p. 58).

Nevertheless, it is important to note that aptitude tests, such as the MLAT, CANAL-F, LLAMA, or Hi-LAB still fail to say anything about the language learning process itself. Robinson (2013, p. 2) emphasises this issue by the following:

Learning a language involves different abilities at different stages of *development*. The MLAT and other current aptitude tests don't measure these.
Learning a language takes place in many different situations and classroom contexts. The MLAT and other current aptitude tests are *insensitive* to these.

1.3 Similarities and Differences: MLAT, CANAL-F, LLAMA, and Hi-LAB

Table 1 below presents similarities and differences between the above described language learning aptitude tests by taking the MLAT as the general established standard (Skehan, 2012).

Table 1 below clearly shows that the MLAT covers the less, and even skips Carroll's inductive language learning factor completely. The LLAMA is more widespread, does more or less without grammatical sensitivity, but "adds a receptive interpretation of inductive language ability as well as more focus on working memory. CANAL-F is the broadest battery of all. It focuses on language analysis and memory (although without an overt concern for working memory), downplays sound, but is far more concerned than the other batteries with attentional function and learning. Skehan (2012) claims that "there is considerable scope for comparative validation research". The Hi-LAB apparently expands the other tests by the factor of error processing and also seems to emphasize more on cognitive and perceptual abilities predicting foreign language learning talent. This is in line with previous researchers who claim that working memory capacity may play the greatest role in predicting language learning success as working memory capacity is limited and only a certain amount of information can be processed at a given time.

Table 1 A comparison of different foreign language aptitude batteries

Components	MLAT (1959)	CANAL- F (2000)	LLAMA (2005)	Hi-LAB (2014)
Phonemic coding ability	x		X	x
Grammatical sensitivity	x			x
Inductive language learning		x	X	x
Paired associates	x	x	X	x
Working memory			X	x
Attentional processing		x		x
WM to LTM connections		x		x
Processing feedback				x

Based on Skehan (2012)

Individuals differ in that amount, and consequently they are either gifted, or not (Lightbrown & Spada, 2013, p. 81). Although this may be true, Skehan (1989) reminds us "that successful learners don't have to be strong in all of the components of aptitude […]", and their "[…] strengths and weaknesses in these different aptitude components may account for their ability to succeed in different types of instructional programmes" (Skehan, 1989 in Lightbrown & Spada, 2013, p. 81).

By and large, the generally accepted components of language learning aptitude tests seem to be the "ability to identify and memorize new sounds, understand the function of particular words in sentences, figure out grammatical rules from language samples, remember new words, and working memory" (Lightbrown & Spada, 2013, p. 80). If aptitude tests can, in fact, be compared to each other is investigated below.

Given that language aptitude tests have been developed to test an individual's talent for language learning, and presupposing that either a person has got talent or not, it may be assumed that two commonly known language aptitude tests – the MLAT and the LLAMA which appears to only complement the MLAT rather than replace it – will produce significantly similar results within the same tested individuals. Yet, this study aims at revealing that language aptitude is a highly complex, multi-componential, and non-unitary ability which in fact cannot be tested that clearly and easily. Types of aptitude measured by different aptitude batteries may not be exactly the same and therefore may be measured differently. Thus, the study aimed at answering following research questions:

1. Will all participants obtain two similar test results?
2. Will there be any significant correlations between the presupposed overlapping test sub-components:
   (a) MLAT III and LLAMA E and D: testing phonetic ability
   (b) MLAT IV and LLAMA F: testing grammatical sensitivity and grammatical inferencing ability
   (c) MLAT V and LLAMA B: testing the memory component of language learning
3. Will there be any further significant correlations?

2 Methodology

2.1 Participants

Thirty-three people (age range 24.8 to 26) were tested for their language aptitude. None of the participants studied linguistics or were experts in this particular field. All of the participants were students of primary school teacher training and four of them have been raised bilingually, the other 29 were native speakers of German with experience in English.

In this cross-sectional study, the MLAT and the LLAMA aptitude tests were administered to all participants.

2.1.1 The MLAT

The study used three parts of the MLAT: III, IV, and V.

- *MLAT III: Spelling clues (50 items)*: is intended to measure English vocabulary knowledge as well as 'phonetic coding ability'. The test requires the candidates to recognize English words written in a reduced form instead of using the conventional spelling system […] and choose from four options the one that is closest to the word in meaning […].
- *MLAT IV: Words in sentences (45 items)*: is intended to measure 'grammatical sensitivity'. In this test, the candidates are given a pair of English sentences, with the first sentence having one word underlined. The candidates are required to select one word with the same grammatical function as that of the underlined word in the first sentence […].
- *MLAT V: Paired associates (24 items)*: in intended to measure 'rote learning ability for foreign language materials'. The test requires the candidates to memorize the English meanings for a set of given words […] in an unknown language and then to choose the meaning of the given word from the multiple choice options provided […] (Sasaki, 2012, p. 316).

2.1.2 The LLAMA

All parts of the LLAMA battery were used:

- *LLAMA B*: intends to measure the candidates' ability to learn large amounts of vocabulary in a relatively short period of time. The program is based on the original vocabulary learning subtask of Carroll and Sapon (1959) […] This version no longer requires any L1 input, so the test is suitable for use with tests of any L1. […]
- *LLAMA D*: is a sound recognition task. It is a new task that does not appear in the work of Carroll and Sapon (1959). It is designed to test if you can recognise short stretches of spoken language that you were exposed to a short while previously. […] These writers suggest that a key skill in language ability is your ability to recognise patterns, particularly patterns in spoken language. If you can recognise repeated patterns, then you are more likely to be able to recognise words when you hear them for a second time. This helps you to acquire vocabulary. It also helps you to recognise the small variations in endings that many languages use to signal grammatical features.
- *LLAMA E*: is a sound-symbol correspondence task. It presents a set of 22 recorded syllables, along with a transliteration of these syllables in an unfamiliar alphabet. Your task is to work out the relationship between the sounds you hear and the writing system. It is particularly good at picking out learners who were able to dissociate sounds from the way they are normally written in English.
- *LLAMA F*: intends to assess grammatical inference that presents you with sentences in an unknown language, and translations of these sentences in your L1. The task is to work out the grammatical rules that operate in the unknown language. Llama F has been designed with a new interface that requires no L1 input (Meara, 2005).

2.2 Procedure

The participants were asked to complete both tests within one week. They were instructed to work rapidly and to spend 50–60 min on the MLAT only and not more than 30 min on the LLAMA.

3 Results

For detecting significant correlations, the Wilcoxon test and Spearman's rank correlation coefficient analysis were applied. Before, an overview over the study's most important descriptive statistics is given in Table 2.

The Wilcoxon test showed a significance level $p = 0.05$. Consequently, the null-hypothesis could be rejected and it can be followed, that the MLAT and the LLAMA produce significantly different results within the same tested individuals. Two different language aptitude tests obviously do not result in equivalent data.

Furthermore, correlations between the presupposed overlapping test categories of the MLAT and the LLAMA were sought by applying Spearman's rank correlation. It could be revealed the MLAT III correlated with the LLAMA E $(r = .40)$. Furthermore, the LLAMA E and the MLAT IV $(r = .369)$ and LLAMA E and MLAT V $(r = .359)$ correlated significantly. No significant correlations between the other previously suggested overlapping test-components MLAT V and LLAMA B, MLAT IV and LLAMA F, and MLAT III could be detected, however. Moreover, a statistically significant correlation was recognized between the LLAMA B $(r = -.535)$, the MLAT V $(r = -.374)$ and the variable age. Moreover, participants scored higher on the MLAT than on the LLAMA and that the LLAMA shows a greater variance than the MLAT. Table 3 illustrates the most relevant correlations between the subtests of the MLAT and the LLAMA and other variables.

Table 2 Descriptive statistics

	M	MD	SD	Variance	Min	Max
Age	24,82	22,00	6,64	44,091	19	46
MLAT 3	23,85	24,00	7,583	57,508	11	39
MLAT 4	20,52	20,00	6,379	40,695	10	32
MLAT 5	17,91	19,00	5,052	25,523	4	24
LLAMA B	56,58	60,00	22,247	494,939	3	85
LLAMA D	31,67	30,00	18,400	338,542	0	75
LLAMA E	76,06	90,00	23,841	568,371	20	100
LLAMA F	50,15	50,00	25,140	632,008	0	90

Table 3 Spearman correlations between variables in the study

			MLAT 3 (50)	MLAT 4 (45)	MLAT 5 (24)	LLAMA B	LLAMA D	LLAMA E	LLAMA F	Age
Spearman's rho	MLAT 3 (50)	Correlation Coefficient	1.000	.613**	−.045	−.042	.277	.400*	.214	.154
		Sig. (1-tailed)	.	.000	.402	.408	.059	.010	.115	.196
		N	33	33	33	33	33	33	33	33
	MLAT 4 (45)	Correlation Coefficient	.613**	1.000	.290	−.051	.072	.369*	.077	−.040
		Sig. (1-tailed)	.000	.	.051	.390	.346	.017	.336	.412
		N	33	33	33	33	33	33	33	33
	MLAT 5 (24)	Correlation Coefficient	−.045	.290	1.000	.188	.006	.359*	.139	−.374*
		Sig. (1-tailed)	.402	.051	.	.148	.487	.020	.220	.016
		N	33	33	33	33	33	33	33	33
	LLAMA B	Correlation Coefficient	−.042	−.051	.188	1.000	.294*	−.006	.309*	−.535**
		Sig. (1-tailed)	.408	.390	.148	.	.048	.487	.040	.001
		N	33	33	33	33	33	33	33	33
	LLAMA D	Correlation Coefficient	.277	.072	.006	.294	1.000	.251	.210	−.049
		Sig. (1-tailed)	.059	.346	.487	.048	.	.079	.121	.393
		N	33	33	33	33	33	33	33	33
	LLAMA E	Correlation Coefficient	.400*	.369*	.359*	−.006	.251	1.000	.102	−.116
		Sig. (1-tailed)	.010	.017	.020	.487	.079	.	.286	.260
		N	33	33	33	33	33	33	33	33

	LLAMA F	Correlation Coefficient	.214	.077	.139	.309*	.210	.102	1.000	−.223
		Sig. (1-tailed)	.115	.336	.220	.040	.121	.286	.	.106
		N	33	33	33	33	33	33	33	33
	Age	Correlation Coefficient	.154	−.040	−.374*	−.535**	−.049	−.116	−.223	1.000
		Sig. (1-tailed)	.196	.412	.016	.001	.393	.260	.106	.
		N	33	33	33	33	33	33	33	33

*Correlation is significant at the 0.05 level (1-tailed)
**Correlation is significant at the 0.01 level (1-tailed)

4 Discussion

The study addressed two major research questions:

(a) will all participants obtain two similar results on both tests?
(b) will there be significant correlations between the presupposed overlapping test categories of the MLAT and the LLAMA?

Research question 1: Will all participants obtain two similar test results?

No, the MLAT and the LLAMA show significantly different results across the tested individuals. The study revealed that language learning aptitude cannot be measured that clearly and easily as two different tests- which aim to test one and the same ability- produce different results. Foreign language learning aptitude should, therefore, be seen as a highly componential ability, which means that some people will score higher in memory and others is analytic abilities, and still others will be better in phonetic coding abilities. This puts emphasis on language aptitude as a multi-componential concept and thus shows that measuring language talent/aptitude is a complex endeavour and should be revisited and rethought. Even though the LLAMA test is based on the MLAT test, both the differ remarkably in terms of results. A reason for this might be that "language aptitude lacks an appropriate definition supported by theories of language learning" (Dörnyei, 2014) and while some focus on the potential of individuals (e.g. Christiner & Reiterer, 2013, 2015) others work on analytical talent and the like.

The results of this investigation questions whether different aptitude test can be taken for all purposes, and thus rather seems to support that researchers need to be highly selective in what form of aptitude they want to work on.

The LLAMA test might be more elaborate for isolating general language aptitude because it does not require any first language, thus it does not test certain language knowledge which can be considered as a great advantage when testing foreign language talent and might lead to more representative results. According to Meara, this test is based on suggestions made by researchers that "a key skill in language ability is your ability to recognise patterns, particularly patterns in spoken language" (Meara, 2005, p. 8). The LLAMA, therefore, seems to be more distinctive which might have a positive impact on the result's representativeness; it is much shorter than the MLAT which consists of many similar items in each test category. The length of the test may have a negative influence on the test-takers and therefore lead to biased results. Finally, it focuses more on working memory capacity than the MLAT. It is commonly agreed by now, that working memory plays an important role in foreign language learning (Lightbrown & Spada, 2013, p. 81; Christiner & Reiterer, 2013, p. 9). The LLAMA thus measures more the raw potential of someone, however, it does not say anything about achievement.

These findings indicate that the LLAMA is more suitable if one seeks to test foreign language learning talent.

Research question 2: Will there be significant correlations between the presupposed overlapping test categories of the MLAT and the LLAMA?

Yes, there is a positive correlation between the LLAMA E and the MLAT III which both test for phonetic coding ability in the broader sense. It can thus be claimed that these test components are rather clear in its intention and goal. Furthermore, there are positive correlations between the MLAT IV and V and the LLAMA E. MLAT IV is supposed to test grammatical sensitivity, MLAT V is supposed to test for memorizing a large amount of words in a very short space of time whereas LLAMA E intends to work out the relationship between sounds you hear and the writing system. Since this ability might overlap with grammatical sensitivity as well as working memory a correlation is not surprising. No significant correlations between the previously suggested overlapping test-components MLAT V and LLAMA B and MLAT IV and LLAMA F could be revealed. It might thus be concluded that the test components are not much related to each other and apparently test different abilities of talent.

The study aimed at demonstrating that language talent seems to be a highly distinct, not unitary and multi-componential phenomenon which consists of several independent cognitive abilities that apparently can only be broadly classified. Different language learning aptitude tests show different outcomes and individual differences even occur among the gifted people. Hence, language aptitude tests seem to test different components of talent and so cannot be compared that simply to each other. It thus may be relevant to define the actual aptitude of interest before testing an individual in order to receive valid and plausible results. Therefore, the general reliability of an aptitude test should be considered carefully since it most likely needs further investigations and a broader test battery to reliably test a person's language talent. Moreover, the practical use of a language learning aptitude test, particularly for the admission to a job or university, should be very wisely and carefully handled as the test outcome, thus aptitude information, might actually do one fairly wrong.

References

Carroll, J. B., & Sapon, S. M. (1958). *Modern language aptitude test*. New York: The Psychological Corporation.

Christiner, M., & Reiterer, S. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology, 4*(874). https://doi.org/10.3389/fpsyg.2013.00874

Christiner, M., & Reiterer, S. (2015). A Mozart in not a Pavarotti: Singers outperform instrumentalists on foreign accent imitation. *Frontiers in Psychology, 9*(482). https://doi.org/10.3389/fnhum.2015.00482

Dörnyei, Z. (2014). Motivation in second language learning. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (pp. 518–531). Boston: National Geographic Learning/Cengage Learning.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York: Routledge.

Doughty, C. (2013). Predicting high level proficiency. *CASL Research fact sheet. April 2013*. Retrieved from: http://www.casl.umd.edu/sites/default/files/1003_FS_April2013.pdf.

Doughty, C. (2014). *Hi-LAB Language Aptitude Profile*. Retrieved from: http://www.casl.umd.edu/sites/default/files/SampleAptitudeProfile_current_Style.pdf.

Ellis, R. (2008). *The study of second language acquisition*. Oxford, UK: Oxford University Press.

Grigorenko, E. L., Sternberg, R. J., & Ehrman, M. E. (2000). A theory-based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *The Modern Language Journal, 84*(3), 391–405. https://doi.org/10.1111/0026-7902.00076

Jørgensen, H. (2008). *What is a talent, how is it identified, and how is it developed?*. Paper presented at the conference "Precollege education and talent development", European Association of Conservatoires (AEC, Oslo May 30, 2008).

Lightbrown, P. M., & Spada, N. (2013). *How languages are learned*. Oxford, UK: Oxford University Press.

Meara, P. M. (2005). *Llama language aptitude tests*. Swansea, UK: Lognostics Retrieved from: http://www.lognostics.co.uk/tools/llama/llama_manual.pdf

Nardo, D., & Reiterer, S. (2009). Musicality and phonetic language aptitude. In G. Dogil & S. M. Reiterer (Eds.), *Language talent and brain activity*. Berlin, Germany: Moton de Gruyter.

Robinson, P. (2013). Aptitude in second language acquisition. In C. A. Chapelle (Ed.), *Encyclopedia of applied linguistics: Language learning and teaching* (pp. 129–133). Oxford: Wiley-Blackwell.

Rysiewicz, J. (2003). Language aptitude testing. Evaluation of two tests' potential to predict foreign language learning success. *12th annual PASE conference.* Retrieved from: https://www.academia.edu/388693/Language_aptitude_testing_Evaluation_of_two_tests_potential_to_predict_foreign_language_learning_success.

Sasaki, M. (2012). The modern language aptitude test (paper-and-pencil version). *Language Testing, 29*(2), 315–321. https://doi.org/10.1177/0265532211434015

Skehan, P. (1989). *Individual differences in language learning*. London: Edward Arnold.

Skehan, P. (2012). Language aptitude. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 381–395). New York: Routledge.

Thompson, A. S. (2013). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *The Modern Language Journal, 97*(3), 685–702. https://doi.org/10.1111/j.1540-4781.2013.12034.x

Part V
Language Aptitude for Pronunciation

Factors Affecting the Pronunciation Abilities of Adult Learners of English. A Longitudinal Group Study

Karin Richter

Abstract This paper seeks to shed light on the question as to why some L2 learners are more successful when it comes to pronunciation mastery than others. The most frequently given explanation for this much debated phenomenon is age and the impact of a critical period for language learning. While there seems to be little doubt about the potential correlation between the age of the learner and the ultimate level of his/her pronunciation mastery, there is more scholarly dispute regarding the question of whether age is the single most important reason for incomplete acquisition. This paper reports on the selected findings of a longitudinal group study which was carried out between 2011 and 2014. In this project, the changes in the perceived degree of foreign accent of a group of adult university students (N = 55) were tracked over the entire duration of their Bachelor studies. The learners were recorded twice, once at the beginning and then again at the end of their studies, reading a text and narrating a picture story. In addition, questionnaires were designed to explore the impact of individual factors that may have played a role. The statistical analysis of the data obtained revealed that the overwhelming majority of the learners managed to ameliorate their foreign accent, yet no single factor could be identified as the most influential driver of pronunciation learning. Instead, a number of variables, most notably motivation, a lack of language anxiety, musicality, and increased exposure to the target language were found to be crucial aspects augmenting pronunciation aptitude.

1 Introduction

One of the most probing issues that has been lying at the heart of Second Language Acquisition (SLA) research is to unravel the puzzle of why some learners are more successful than others. This phenomenon is often epitomized by the notion of

K. Richter (✉)
Department of English and American Studies, University of Vienna, Vienna, Austria
e-mail: karin.richter@univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_18

foreign language aptitude with its underlying assumption that "there is a specific talent for learning foreign languages which exhibits considerable variation between individual learners" (Dörnyei & Skehan, 2003, p. 590). Particularly interesting in this context seems to be the field of phonological competence, where the amount of variation that can be observed among mature L2 learners appears to be marked by a great scarcity of native-like control. In fact, it has been claimed that merely 5%–15% of adolescent L2 learners obtain native-like or "accent-free" speech (e.g. Birdsong, 2005).

Research has suggested that foreign accent variation is subject to the workings of psychological (e.g. attitude and motivation), socio-cultural (e.g. exposure time, teaching method) and – above all – biological (e.g. age or gender) factors (Reiterer, 2009) which leave their imprint on second language learning aptitude. In the literature, the most frequently mentioned variable accounting for L2 mastery is age and the influence of the Critical Period Hypothesis (CPH), according to which complete L2 mastery will be severely hampered once the learner has passed a certain maturational stage in his/her life (Lenneberg, 1967). More recently, however, attention has shifted away from biological constraints to the impact of psychological and socio-cultural factors on the premise that age can hardly be the only factor that plays a role in predicting a learner's attainment in the L2. Instead, a number of other non-biological individual differences have proven to be pivotal in shaping the path of the learner's development (e.g. Moyer, 2013).

Drawing on the most recent findings in the field, this paper will first of all provide a brief overview of the most frequently mentioned factors affecting pronunciation aptitude. Then, it will proceed to report some of the findings of my unpublished PhD thesis[1] (Richter, 2015a), which aimed to identify the driving forces in the phonological development of a group of adult Austrian learners of English.

1.1 Review of the Literature: Individual Factors Affecting Pronunciation Abilities

The relentless quest for predictable patterns of language development and factors influencing the performance of a learner have guided SLA research for decades. A closer look at empirical studies investigating those individual factors that have been claimed to affect pronunciation mastery gives rise to the assumption that there is a high degree of ambiguity and complexity involved in focusing on these variables in isolation. Indeed, there appears to be a constant interplay of a number of these learner propensities contributing to the development of an individual's L2 pronunciation. Furthermore, it should be noted that these studies vary greatly in terms of their design and methodology (subjects studied, elicitation and rating techniques used, stimuli chosen etc.). These methodological differences, as Piske, MacKay,

[1] Some sections of this paper have been taken verbatim from my unpublished PhD thesis (Richter, 2015a).

and Flege (2001) point out in their review, can – to a certain extent – be held responsible for the often conflicting results that have been yielded.

In the following, the variables which are most commonly associated with L2 phonological development, namely age, gender, musicality, attitude and identity, motivation, anxiety, formal pronunciation instruction, and exposure to the target language will be explored on the basis of the most recent research findings. Although this section discusses these factors individually, this needs to be seen as an organizational convention that does not deny their inherent interconnectedness as various references to this phenomenon will be made throughout.

1.1.1 Age

There has been a long-standing debate in second language acquisition regarding the extent to which the ability to acquire a foreign language can and should be linked to age and maturational constraints of the aging brain. As proficient mastery of the target language is far less common among adult learners (e.g. Han, 2004; Nakuma, 1998; Scovel, 2000; Selinker & Lakshmanan, 1992), the question as to whether or not some kind of critical period for language learning exists has been central to SLA research. According to the Critical Period Hypothesis there is a certain window in the human developmental process when the ability to learn a new language reaches its peak. Thus, if the learner is exposed to new input during this window, theorists believe that it is certain that he or she can become proficient, but once this window closes, the chances for mastery fade. Among those who find a strong relationship between the age of exposure and ultimate language mastery are for instance Johnson and Newport (1989), Newport (1990) or Patkowski (1994).

Some researchers have even gone one step further by postulating that a critical period exists only in the realm of pronunciation. Scovel (1988), for example, asserts that – in contrast to other language skills – "phonological production is the only aspect of language performance that has a neuromuscular basis" (p. 101). Indeed, adults seem to vary greatly in their L2 pronunciation aptitude – both with regard to segmental as well as suprasegmental parameters of spoken language (e.g. Golestani & Zatorre, 2009; Jilka, 2009). However, a number of scholars have cautioned that the variable age should not be looked at in isolation. Along these lines, Moyer (2007), for instance, purports that psychological factors are in fact more important than biological maturation claiming that age could be an indirect rather than a direct factor in as far as it has to be seen in correspondence with other – biological, psychological and environmental – variables.

1.1.2 Gender

Intuitively, gender seems to be a very appealing variable in SLA. After all, the commonly-held folk wisdom that girls are better language learners than boys still exists. Yet research findings are not conclusive regarding female advantage in foreign language learning in general (e.g. Brantmeier, Schueller, Wilde & Kinginger, 2007) and pronunciation learning in particular (e.g. Elliot, 1995; Purcell & Suter, 1980). It should be noted, however, that very few studies solely focus on gender and pronunciation. Instead, gender is often combined with other variables such as attitude, motivation, age or exposure.

Empirical studies in this field dating back to the late twentieth century (e.g. Elliot, 1995; Purcell & Suter, 1980) did not identify gender as a reliable predictor of degree of L2 foreign accent, whereas more recent investigations have (e.g. Dogil & Reiterer, 2009; Khamkhien, 2010). Jahandar, Hodabandehlou, Seyedi & Mousavi Dolat Abadi (2012), for example, point out that although female students often outperform their male colleagues in producing accurate consonants (but not vowels), this does not give rise to assume complete superiority of female over male subjects.

In another experiment carried out in Germany on L2 pronunciation talent, Reiterer et al. (2011) noticed a significant gender difference in one of their tasks, namely the imitation of Hindi sounds, where the scores for the male imitators were higher. They speculate that the reason for this may lie in the task type requiring a speech imitation skill which was devoid of syntactic and semantic operations. Apparently, when it comes to motor skill learning, recent evidence (Dorfberger, Adi-Japha, & Karni, 2009) shows that male learners have a major advantage over their female counterparts. Giftedness research in general suggests that gender differences are greater in talented than in average ability individuals (Preckel, Goetz, Pekrun & Kleine, 2008). They attribute this to the fact that evolutionary theories predominately consider males as more located in the extremes of the normal distribution curve, whereas females tend to be more represented towards the mean (with respect to any kind of ability).

1.1.3 Musicality

It is commonly believed that people who are skilled at music have a sensitive ear and can discriminate between sounds more accurately and consequently imitate speech sounds better. This means that musically trained individuals have an advantage over non-musicians in as far as they pick up various aspects of an L2, in particular the pronunciation of L2 sounds, more easily. Hence the question which scholars from across academic fields have tried to answer is whether there is a link between musical ability and language proficiency.

Previous scientific findings are – again – inconsistent. Whereas some researchers have identified a positive correlation (e.g. Nakata, 2002; Tanaka & Nakamura, 2004), others have not (Flege, Munro & MacKay, 1995; Flege, Yeni-Komshian & Liu, 1999). However, more recent research into musical aptitude and its influence on L2 pronunciation mastery draws a somewhat clearer picture. Studies conducted by Pastuszek-Lipińska (2004), Slevc and Miyake (2006), Milovanov, Huotilainen, Valimaki, Esquef, and Tervaniemi (2008), Nardo and Reiterer (2009) or Christiner and Reiterer (2013) postulate that musicians (both vocalists and instrumentalists) generally find it easier to acquire new sounds, and see this as evidence for the link between phonology and music.

Very recently, evidence from neuroscientific research has shown that musicians exhibit specific differences in their brain structure when compared to non-musicians (e.g. Seither-Preisler, Parncutt & Schneider, 2014). It has been claimed that these biological alterations may in fact lead to an improvement of both, music and speech perception on the one hand (e.g. Oechslin, Meyer & Jäncke., 2010), but also to enhanced literacy and attentional skills (cf. Seither-Preisler et al., 2014), thereby suggesting a distinct link between musical talent and language talent.

1.1.4 Attitude and Identity

One of the most frequently mentioned predictors of the acquisition of L2 pronunciation is the learner's attitude towards the target language and culture. As a matter of fact, the – conscious or unconscious – decision as to whether or not to adopt or imitate a particular accent can often be traced back to the speaker's relation to the target language and community. Along these lines, Kenworthy (1987) notes that

> [s]ome individuals seem to be impervious and even after a long time will absorb only some turns of phrase and the pronunciation of a few individual words. Others seem very receptive and begin to change their accent almost as soon as they step off the plane! (Kenworthy, 1987, p. 7)

Given that identity and the concept of self are driving forces for the successful acquisition of L2 phonology, concretizing this accent-identity link has proven to be a challenging undertaking. In this respect, Moyer (2004) found that some advanced learners enjoyed taking on a particular accent when travelling abroad in order to be passed for a native speaker of the L2. On the other hand, some learners may hold on to their foreign accent to ascertain the link to their linguistic origin (Moyer, 2004).

1.1.5 Motivation

Another crucial ID variable which calls for closer scrutiny is motivation. Moyer (2004) regards the learner's motivation as "a construct that uniquely represents many orientations simultaneously: conscious effort, intentionality, and planning towards a specific goal" (Moyer, 2004, p. 39). Accordingly, it is based on "interest or curiosity to know more, along with perceived likelihood of success and reward" (Moyer, 2004, p. 39) and has been named the strongest and most influential factor determining the success or failure of learning a second language (Dörnyei, 2010).

Although the importance of motivation in learning a second language has been investigated by a great number of researchers (e.g. Dörnyei, 2010; Gardner & Tremblay, 1994; Smit, 2002; Smit & Dalton, 2000), there seems to be a paucity of research into students' motivation to improve their speaking skills. Despite this general lack of attention, studies into the interrelation between motivation and pronunciation achievement yield a fairly consistent and clear picture. Conceivably, having a personal or professional goal for learning English can influence the need and desire for native-like pronunciation (e.g. Bernaus, Masgoret, Gardner & Reyes, 2004; Gatbonton, Trofimovich & Magid, 2005; Moyer, 1999). In the concluding remarks of their review of research on adult acquisition of English, Marinova-Todd, Marshall, and Snow (2000) assert that adults can become highly proficient – even native-like – speakers of second languages, if motivated to do so. What most of these studies have in common is a fairly unanimous suggestion that learners with a high degree of motivation as regards the foreign language in general and pronunciation mastery in particular are likely to achieve better results than those whose motivation is low.

1.1.6 Anxiety

A further factor that merits examination and is in fact closely linked to motivation is the notion of foreign language anxiety (LA). Scholars have found that the foreign language learning process is particularly prone to arousing anxiety (e.g. Campbell & Ortiz, 1991; MacIntyre & Gardner, 1989; Reid, 1995). Liu and Huang (2011) or Olivares-Cuhat (2010) have even gone as far as to claim that LA should in fact be considered the strongest and most powerful predictor of success in FL learning.

In line with the observation that language anxiety seems to have a negative influence on foreign language mastery in general, studies investigating oral production/pronunciation and learner apprehension have found negative correlations (e.g. Liu, 2006; Stephenson Wilson, 2006; Woodrow, 2006). For instance, learners are sometimes reported to experience anxiety when giving presentations in class, interacting with a native speaker, or being corrected while speaking (Mak, 2011). Anxious foreign language learners also mention challenges that can be directly linked to pronunciation. By way of illustration, they complain about difficulties discriminating the sounds and feel embarrassed because of their pronunciation errors (Price, 1991). Similarly, Ohata (2005) points out that unrealistic beliefs can cause greater anxiety and frustration, particularly when beliefs and reality do not match. He asserts that if the learners believe that pronunciation is the single most important aspect of language learning, it comes as little surprise that they will be frustrated to find the reality of their poor speech pronunciation even after learning and practicing for a long time.

1.1.7 Formal Pronunciation Instruction

It appears that only very few studies so far have looked into the effectiveness of explicit pronunciation instruction (e.g. Bongaerts, Van Summeren, Planken & Schils, 1997; Derwing & Munro, 2005; Moyer, 1999) despite its obvious importance in successful communication. Generally speaking, divergent results have been found as to the effects of formal instruction on L2 pronunciation mastery. It seems that the quality of input (e.g. specific phonetic training) is of greater importance than the quantity (i.e. time spent studying the pronunciation of the target language). Taken all together, the studies on the amount and type of input have not provided enough convincing evidence upon which variance in L2 phonological acquisition might be solely based. In spite of the finding that late L2 learners can benefit from formal phonological instruction, the input they receive should not be looked at in isolation. Rather, other factors such as attitude, motivation, age of learning, length of residence, etc. are likely players, too.

1.1.8 Exposure to the Target Language

Recent socio-political, economic, and educational changes support the general assumption that the importance of the variable exposure to English is currently gaining momentum in language learning. The promotion of seemingly borderless student mobility, the ready availability of native-speaker English on the worldwide web and also the current trend of educational institutions to teach content courses in English have dramatically increased the amount of English learners from all over the world are nowadays exposed to. This additional engagement – and it is often engagement rather than merely passive exposure – with the L2 can be assumed to leave an imprint on the learners' foreign accent. To take account of these changes, the factor exposure will be discussed here as three distinct sub-variables, namely study abroad (SA), media exposure (ME), and English-medium instruction (EMI).

Study Abroad (SA)

In the context of the current social, cultural and educational developments triggered by globalisation and internationalisation endeavours, student exchange programmes have come to assume an important role in L2 learning policies particularly as a means to promote multilingualism (cf. Kinginger, 2009). A widely held assumption in this respect is that students who study abroad return home with substantially increased foreign language competences, particularly in areas such as fluency, pronunciation, or vocabulary.

Research has mostly concurred that study abroad can indeed be a beneficial way of acquiring a foreign language (e.g. Freed, Segalowitz & Dewey, 2004; Lafford, 2006; Lafford & Collentine, 2006). While a great number of investigations into the effect of SA on L2 acquisition have been carried out in the field of selected morpho-syntactic areas (e.g. Howard, 2005; Isabelli & Nishida, 2005), less work is available discussing the specific effects of SA on L2 pronunciation.

With a focus on the pronunciation of SA vis-à-vis at-home (AH) learners, Diaz-Campos (2004, 2006) found that although both groups made progress, there was no consistent advantage for the SA group. Concerning the oral production of native-like variants, however, this study revealed that SA students appeared to generate more target like variants than AH students in informal (conversational) tasks

(Díaz-Campos, 2006, p. 37), whereas no major differences between groups could be found in more formal tasks (e.g. reading aloud). Similarly, Lord (2006) examined students' pronunciation and mimicry abilities in Spanish before and after SA. Again, she found that the learners' overall pronunciation skills did not improve significantly. Another study conducted by Avello, Mora and Perez-Vidal (2012) explored the perception of foreign accent in a SA context by assessing the effect of a 3-month SA programme on the pronunciation of a group of 23 Spanish learners of English. Although they found a slight yet non-significant improvement in perceived foreign accent, a notable decrease was detected in pronunciation accuracy scores after SA.

Media Exposure (ME)

Media exposure to English in the form of cable TV, the internet, computer games, music or radio has also been researched in connection with L2 attainment, although on a much smaller scale. As Uskoski (2011, p. 16) points out, increased access to the internet and the growing influence of the global community have strongly influenced language learning since the 1990s. Owing to technological advances, access to the English language outside the classroom is rapidly spreading all over Europe. According to Livingstone (2002) "the media today operate as pervasive, yet often imperceptible, elements in the everyday cultures of children and young people" (p. 286), She purports that the media can and do have a favourable effect on students learning English as an L2, due to the fact that a great number of media genres in Europe are now available in English. In the context of SLA research, it comes as little surprise that ME studies have discovered that receptive skills (i.e. listening and reading) benefit most, whereas productive skills (i.e. speaking and writing) are less affected (e.g. Pickard, 1996).

English-Medium Instruction (EMI)

In the twenty-first century, the internationalization of education and the wish to compete globally have caused an increase in English-medium instruction (EMI), the teaching of content classes in English. This international call for EMI is deeply rooted in the common belief that language learning takes place incidentally during content delivery in a foreign language. And indeed, empirical research in the field of language learning and teaching at secondary school level seems to support the assumption that a second language is learned most effectively when it is used to convey content that is both interesting and relevant to the learner. However, to date tertiary education has seen little empirical evidence that this is in fact the case.

As far as pronunciation is concerned the impact of English-medium instruction seems to be altogether more moderate. Indeed, reviewing the latest and most relevant studies related to the effects of English-medium teaching on language competence, it appears that spoken production in general and pronunciation in particular have only received minor attention, especially when focusing on the tertiary level of education. Nevertheless, a number of researchers from Spain have looked at pronunciation and foreign accent in secondary school settings, where the use of English to convey content is commonly referred to as CLIL (Content and Language Integrated Learning). What these empirical investigations conducted by Ruiz de Zarobe and Jimenez Catalan (2008), Gallardo del Puerto, Garcia Lecumberri and Gomez Lacabex, (2009) or Rallo Fabra and Juan Garau (2010) seem to have in common is that they generally fail to observe a positive impact of CLIL on the learners' pronunciation skills. The authors attribute their findings to a lack of sufficient native-speaker input since the teachers in their programmes tend to be native speakers of Spanish rather than English.

Conceivably, pronunciation is one of the skills that seem to remain largely unaffected by EMI. The main reasons for this phenomenon are threefold: quantity of input (only few courses are taught in English), quality of input (referring to the teachers' own pronunciation skills) and a lack of long-term observation.

1.2 Research Questions

As discussed above, we are faced today with an abundance of views on ID variables that impact on second language learning aptitude. Although some of the driving forces linked to phonological development are indeed likely to affect L2 pronunciation achievement (most notably age, motivation and language learning anxiety), divergent results have been found in empirical investigations so far.

Against the backdrop of a general scarcity of longitudinal data in the field of pronunciation learning, this group study attempts to explore a number of these potential factors and their importance in relation to the development of adult learners' pronunciation abilities over an extended period of time. Thereby, two interrelated questions will be addressed:

1. How does the adolescent learners' degree of foreign accent change in the course of 3 years of tertiary education?
2. Which individual variables can be identified as the driving forces in the students' phonological development?

Whereas the first question is intended to challenge the commonly held belief that the L2 pronunciation learning process of adult students is severely hampered by age-related constraints, the second question looks at the broader picture by taking other individual factors into account, thereby seeking to identify the strongest predictor of phonological change in the learners.

1.3 Hypotheses

Drawing on the tenets of the Critical Period Hypothesis and the literature on factors influencing pronunciation mastery as reviewed above, it can be hypothesized that no matter how distinctively marked the individual characteristics of each student are, the adult learners in the present project are likely to experience only minor if any positive development with respect to their foreign accent. Nevertheless, it is viable to speculate at this point on the profile that emerges of an English language learner who has the potential to develop native-like pronunciation: Conceivably, the student began to learn English in early childhood (1), is probably female (2), plays a musical instrument or enjoys singing (3), has a positive attitude towards the English language and culture and is willing to become part of this particular L2 community (4), is highly motivated to learn the language and improve his/her pronunciation and also cares about his/her own foreign accent (5), is not afraid to speak English (6), has attended explicit pronunciation classes (7), and has been exposed to a considerable amount of L1 English both inside and outside the classroom (8) (cf. Richter, 2017).

2 Methodology

2.1 Participants

The learners (N = 55) who participated in this study were all undergraduate Business students at the UAS (University of Applied Sciences) Vienna. The subjects' phonological development was tracked over the entire duration of their studies (from 2011 to 2014). All the participants had a fairly similar profile in terms of age range (i.e. 19–25), L1 (Austrian German), English language proficiency (B2) and average age of onset of learning (8 years). To take account of the increasing importance of in-class exposure to English in the form of EMI, students from two different cohorts starting in the winter term 2011 were asked to participate in this study. Therefore, roughly half of the students (N = 25) were enrolled in the university's English-medium programme, which means that they had up to 50% of their classes taught in English by native speakers and half of the learners (N = 30) came from the regular German programme with one two-hour ESP class per week.

2.2 Instruments and Procedures

In order to address the research questions outlined above, the students were recorded twice, once at the beginning of their studies (T1 = 2011) and then again at the end (T2 = 2014) in the university's soundproof radio studio. Two different tasks were

Table 1 Statistics of the development of foreign accent

Task	Mean (SD)			t-test	
	T1	T2	Average difference	t	p
Reading	4.77 (± 1.50)	5.62 (± 2.03)	+0.85	−5.53	***
Speaking	5.22 (± 1.51)	6.03 (± 2.13)	+0.81	−5.08	***
Average	4.99 (± 1.44)	5.83 (± 2.01)	+0.83	−6.01	***

*Significant at the 0.05 probability level
**Significant at the 0.01 probability level
***Significant at the 0.001 probability level

assigned to them, namely the reading of the standard IPA text "The North Wind and the Sun" and the narration of a Gary Larson cartoon (cf. Richter, 2015b). While the reading task was included to facilitate a controlled coverage of the phonemic inventory, the quasi-spontaneous picture story allowed for a more natural type of speech.

When the recording process was completed, each sound file was rated by seven carefully chosen pronunciation experts. These raters had more than a decade of experience teaching pronunciation classes and assessing their students' pronunciation skills at the University of Vienna. In order to capture a wide range of perspectives, three male and six female lecturers stemming from various linguistic backgrounds (L1 speakers of American English, British English and Austrian German) were included. In their courses, these teachers had always examined their (predominantly Austrian) students' pronunciation skills in pairs. This means that they were not only distinctly familiar with the peculiarities of the Austrian accent in English but they also shared a common understanding of grading foreign-accented speech.

Since there is no standardized tool to measure foreign-accented speech (e.g. Piske et al., 2001), it was decided to replicate the evaluation tool designed by Dogil and Reiterer (2009) for their exploration of the multiple manifestations of pronunciation talent. Hence the raters used a web-based rating tool comprising a visual analogue scale to classify the sound files on a gradation spectrum ranging from "very strong foreign accent" (0) to "no foreign accent" (10).

To elicit biographical data from the participants regarding individual factors that are frequently linked to the level of pronunciation mastery in second language acquisition, two questionnaires were designed; the first one (Q1) was administered to the students at the beginning of their studies (in 2011) and the second one (Q2) at the end (in 2014). These questionnaires were again largely based on those administered by Dogil and Reiterer (2009).

3 Results

One of the key aims of this project was to track the students' phonological development over an extended period of time. Table 1 presents the statistical analysis of the scores obtained from the raters for all the 55 learners at T1 and T2 respectively.

Table 2 Factors influencing pronunciation development

Code	Variable	p	Partial eta squared
MOT	Motivation to improve pronunciation	0.018	0.964
	Motivation to improve English language skills	0.012	0.976
ANX	Anxiety	0.022	0.956
ATT	Attitude towards English	0.442	0.311
GEN	Gender	0.115	0.784
MUS	Musicality (singing)	0.026	0.991
EXP	English-medium instruction	0.019	0.963
	Media exposure	0.128	0.761

A t-test for dependent samples performed on the data revealed that for both tasks (reading and speaking) a significant improvement in the development of the learners' pronunciation skills could be detected with $p < 0.01$. In absolute figures, this means that from all the 55 subjects in the present study, the overwhelmingly majority, namely four fifth ($N = 32$) were found to have ameliorated their foreign accent.

Interestingly, both skills (reading and speaking) seem to have developed similarly with the score for reading showing an increase of 0.85 and the narration of the picture story (i.e. speaking) a rise of 0.81 which results in an average development of +0.83.

Trying to identify the strongest predictor for the changes in the perceived degree of foreign-accentedness, the biographical data collected in the questionnaires were matched with the average difference calculated for pronunciation development. Therefore, a multi-factorial one-way analysis of variance (ANOVA) with the difference in the development as the dependent variable was conducted. The statistical analysis revealed the following results (Table 2).

What can be deduced from the table is that no single variable was found to be highly significant ($p < 0.001$). The following four parameters, however, were detected to be significant ($p < 0.05$): motivation, lack of language learning anxiety, musicality and exposure in the form of English-medium instruction. To measure the effects of these variables and possible interactions, partial eta squared was calculated to examine effect sizes. In this case, partial eta-squared showed that the influence of the variables was very high with values >0.9. Yet no relevant interaction effects were disclosed.

To facilitate a better understanding of how these variables then impact on the individual level, it was deemed useful to analyse two learner profiles in more detail thereby seeking to reveal indications of the determining factors that shape a high performer and a low performer.

The students whose performances were selected for a more detailed description are on the one hand high-achiever informant #61 with the most remarkable progress regarding the mitigation of the foreign-accentedness of her speech (overall +3.6) and low-achiever informant #49 whose scores showed the lowest development (overall −1.3). In the following, the data gathered in the two questionnaires from both informants will be sketched and correlated with the results obtained in the rating to see if any key ID variables can be identified that may have contributed to the development of the respective learner.

3.1 High Achiever #61

Looking at the scores that the judges assigned to informant #61, it can be seen that the student displays a remarkable improvement from T1 to T2 of M = +3.6 overall (+3.5 for the reading of the text and +3.8. for the narration of the story). Interestingly, the progress in the speaking task was rated higher than that for the reading of the text. This means that the amelioration of the student's foreign accent was considered to be more salient in the quasi-spontaneous speaking task than in the reading.

In order to draw a more comprehensive profile of this particular student's performance, it seems essential to explore those ID variables that may have impacted on the development. According to the information provided in questionnaire 1, informant #61 started to learn English at the age of 6 (1), is female (2), neither plays a musical instrument nor enjoys singing (3) but holds a highly positive attitude towards the English language (4). Also, she displays a distinctively high degree of motivation to learn the language and to improve her pronunciation skills (5). The student claims not to be afraid to speak English (6) and has never received any explicit pronunciation instruction (7). As far as exposure to the target language (8) is concerned, at T1 she reports to have spent a few days in England, 2 weeks in the United States and 3 months working in Malta at a hotel reception as part of her secondary school education. Regarding media exposure, she asserts that she sometimes watches movies or TV programmes in English, reads books or newspapers in English and occasionally visits English websites. Interestingly, this student was not registered in the EMI but in the German programme, which means that she did not benefit from extensive exposure to L1 English in class. However, during her studies, informant #61 did a four-month internship in San Francisco, after which she took a 2 month holiday in Hawaii. So all in all she spent 6 months in an English-speaking country in the last year of her studies. When asked what she considered important factors that contributed to the improvement of her English language skills outside class, she replied that friends and holidays have been highly influential.

3.2 Low Achiever #49

The second learner to be portrayed here is informant #49 whose performance was chosen for closer scrutiny as the scores he received for T1 and T2 showed the most noticeable decline in pronunciation skills. This means that of all the 55 informants, the development of his foreign accent was rated by the listeners as the least favourable of all the participants. His pronunciation was clearly considered to be more heavily accented at T2 than at T1.

Low achiever #49 started to learn English at the age of 10 (1), is male (2) and does not consider himself a musically talented person (3). Overall, the information he provided in questionnaire 1 shows that the student generally holds a very positive attitude (4) towards the English language. In terms of motivation (5), the answers provided create a more diverse picture. For instance, he sees good English language skills as an important goal in his life, but he does not appear to be very interested in indulging further in the English language and culture. What is particularly striking is the fact that he expresses a low level of concern about his own pronunciation. Although he claims that he would like to improve this particular skill, he does not feel the need to sound like a native speaker and definitely does not want others to think that he is in fact one. When considering the variable anxiety, it appears that the student is not afraid of speaking English either in the classroom or outside (6). He has never attended any pronunciation classes in English (7). Regarding exposure to the English language (8), he spent 3 weeks in the US before he started to study at university. In addition, he reports that he never watches films in English, never reads English books or newspapers/magazines, and never listens to English radio stations. Occasionally he watches English TV and visits English websites. Interestingly, the data gathered in Q2 revealed that during his studies, the student spent a semester at a South Korean university where he spoke predominantly English with non-native speakers of the language.

4 Discussion

This study set out to examine factors that impact the pronunciation abilities of adult EFL learners. To this end, a group of adolescent Austrian business students from the University of Applied Sciences Vienna served as informants in a project that endeavoured to measure changes in the perceived degree of foreign accent and to identify potential factors that shape the process of acquiring pronunciation skills in an L2.

In reviewing the most recent and most influential studies on these variables, it was found that in SLA research so far no individual factor has been singled out as the most reliable predictor for phonological change. Although the age of the learner seems to be a promising candidate, research in this field is highly inconclusive. The same appears to be true when trying to measure the impact of other individual factors such as gender, motivation, attitude, musicality or exposure to the target language.

Overall, the data collected in this project provide convincing evidence according to which the existence of a critical period for accent with absolute maturational limits looks less and less plausible. The finding that four fifth of all the 55 adult subjects managed to ameliorate their foreign accent by far surpasses any expectations that other studies could have raised. By all indications, this suggests that in the present instance the biological constraints of the aging brain have been superseded and supplanted by other contextual capacities that have come to the fore and thereby left their unmistakable imprint on the development of the foreign accent. In this respect, the hypothesis made in Chapter 1.3 according to which the adult learners were likely to experience only little if any progress at all needs to be rejected. Obviously, age cannot be regarded as the single most influential factor affecting the pronunciation abilities of adult learners.

When considering the data derived from the analysis of other individual factors which are commonly said to impact on phonological attainment, the results of the present project give rise to the assumption that although no single variable alone can be held responsible for the observed pronunciation development, a number of interwoven variables seem to be fundamental drivers in promoting learning. Taking a closer look at these factors (i.e. motivation, lack of language learning anxiety, musicality, and English-medium instruction), it is interesting to see that of these four variables three are either psychological in nature (i.e. musicality, language learning anxiety, and motivation) and only one, namely exposure to English in the English-medium classroom, can be associated with socio-cultural parameters. One might therefore argue that in this particular case 'psycho(bio)logical' factors (cf. Reiterer, 2009) have prevailed in determining the degree of success of this particular group of language learners.

Arguably the most complex variable in this respect is motivation, which has already been named "the primary impetus to initiate language learning" (Dörnyei, 1998, p. 117). As Dörnyei points out, a high degree of motivation can – to a certain extent – compensate for deficiencies in the learner's language aptitude (Dörnyei, 1998). Closely associated with the variable motivation appears to be the variable language learning anxiety or rather the lack thereof. In fact, a number of researchers have investigated the relationship between FL mastery on the one hand and FL anxiety and motivation on the other hand (e.g. Khodadady & Khajavy, 2013; Liu & Huang, 2011; Tóth, 2007). In their studies they found that anxiety was positively and significantly related to motivation. Therefore it comes as little surprise that the present study confirms the positive influence of these two psychologically driven factors.

In addition to motivation and anxiety, the factor musicality is an interesting determiner to predict a learner's L2 phonological competence as it is often mentioned in the context of language learning aptitude. So far the question whether music and language share common neural resources has not been sufficiently answered. What appears to be crucial in this regard, however, is the potential correlation between musicality and motivation. A study conducted by Seither-Preisler et al. (2014), for instance, confirms this assumption. The authors argue that biological factors (like genes or hormones) promote maturational plasticity with increasing age and thereby foster the development of both perceptual and cognitive skills. They also assert that musicality as a biological predisposition may have a positive effect on practicing behaviour (i.e. motivation). It might thus be assumed that motivation could function as a pre-existing biological variable that affects language talent in a favourable manner.

As far as the factor exposure to the target language is concerned, crucial differences in its impact on the adult learners' pronunciation competence could be detected. A close examination of the sub-variables 'exchange semester abroad', 'internship abroad', 'media exposure' and 'English-medium instruction' has shown that the most influential of these aspects was clearly English-medium instruction. Those students who were enrolled in the English-medium programme were generally found to have progressed more, were more likely to study abroad (84% as opposed to 23%) and also to do their internship abroad (20% and 17% respectively). In terms of media exposure it was revealed that the learners in the EMI programme were considerably more exposed to the English language not only in class but also beyond with the internet being the most frequently mentioned source. The degree to which EMI seems to have accounted for changes in the degree of foreign accent undoubtedly comes as a surprise. So far no empirical investigation has managed to prove phonological benefits of teaching content classes in English. However, it has to be noted that the high quantity (more than 50% of the content classes were taught in English) and quality of input (the majority of the teachers were L1 speakers of English) is rather exceptional in European English-medium classes.

In brief, the data gathered in the present project advance the view that no single factor alone can be held responsible for the observed pronunciation development, rather a combination of psychological and socio-cultural variables, such as motivation, lack of language anxiety, musicality and English-medium instruction seems to have been pivotal in promoting phonological acquisition. Yet, predicting their relative prominence at various stages of the learning process is difficult. This finding can be seen in light of the Dynamic Systems Theory (DST) proposed by Larsen-Freeman (1997), suggesting that language learning as such is an inherently dynamic process that is largely determined by a set of inter-dependent variables that interact over time (cf. De Bot, Lowie & Verspoor, 2007). Applying a DST approach to second language acquisition is said to capture both social and cognitive aspects and thereby shows how their interaction can lead to development (e.g. De Bot et al., 2007; Herdina & Jessner, 2002). This view is largely supported by the results obtained in this project corroborating the notion that language learning in general and the development of foreign accent in particular are distinctly shaped by an interconnected web of individual factors. In their complexity, these factors ostensibly interact with the learning context in a dynamic manner. In other words, socially and psychologically dependent variables are particularly prominent for any learner and are likely to influence the effort the learner puts into trying to sound native-like.

As these individual factors are difficult to generalize, two participants of the study were selected for further investigation, namely high achiever #61 and low achiever #49. These two informants clearly differ regarding a number of variables that are undoubtedly interrelated and combine to create a vague picture of a high-potential (HP) learner and a low-potential (LP) learner. Whereas it could be argued that (slightly) earlier AOL, a positive attitude and higher motivation to be observed in the HP learner have reinforced the experience the student made in the course of an internship in the United States, a later onset of learning together with a lower interest in the English language and culture paired with a lower degree of concern about his own accent resulted in the LP learner's negative development after a stay at a Korean university. However, two of the four parameters (i.e. musicality, motivation, lack of language learning anxiety and EMI) which have been found to have distinctly shaped the language learning process of the entire student population in the present project, cannot be attributed to those two learners: These two students do not see themselves as musically talented and they are not enrolled in the EMI programme.

For the most part, informant #61 fulfils the criteria outlined in the section on profiling a high potential pronunciation learner. The most noticeable deviations, however, concern the fact that she has not received explicit pronunciation instruction and that she neither plays a musical instrument nor enjoys singing. Also, she was not enrolled in the EMI programme. What seems to stand out noticeably in her language learning biography is the fact that she pursued an internship in the US in the course of her studies at the UAS Vienna. Conceivably, the experience of working abroad exceeds the experience of studying abroad in that it makes greater demands on the students. During an internship, the students tend to be involved more deeply in the target culture and they also use the L2 for a greater variety of purposes. Whereas learners in a regular study-abroad context often continue their education in a different country, in international internships students are frequently challenged to grow in three different ways; namely professionally, socio-culturally and linguistically. These placements with companies frequently require a high amount of theoretical knowledge of the subject matter (i.e. business concepts as studied at home) that is put into practice as well as socio-cultural and linguistic competence in the daily interaction with colleagues, suppliers, or customers. These central demands combine to constitute essential prerequisites for coping successfully with both content and language matters. Without doubt, the issue of internships abroad is an intriguing aspect when researching factors influencing foreign language learning that cannot be neglected in future SLA research.

Of course, great care needs to be taken to generalize the claims made here based on the analysis of only two admittedly opposing learner profiles. However, it is fascinating to see that the qualitative analysis of these two learners greatly supports the findings discussed above according to which it is difficult if not impossible to identify the most important factor which is responsible for the success or failure of a language learner. Instead, it is more appropriate to see these ID variables as part of an intricate and complex web of factors that are unique in every single learner.

5 Conclusion

This paper has analysed the impact of individual variables on the pronunciation learning process of adult EFL students. In doing so, fresh insights into a much debated phenomenon have been provided. Hence, the presented results make noteworthy contributions to the current literature on aptitude in SLA in at least two major respects.

First of all, the findings generated in this study contribute further to our understanding of second language acquisition in general and pronunciation learning in particular. By and large, the results presented here overwhelmingly exceed all hopes raised by empirical research so far. The main finding according to which the overwhelming majority of the adult university students in this project experienced a clear amelioration of their foreign accent undeniably rejects the widely held view that language learning development is to a large extent underpinned by special bioprogramming stemming from a general loss of plasticity in the aging brain. In fact, the opposite seems to be true and substantial gains could be revealed. Even though these results essentially have to be seen within the scope of the given setting, the study also offers valuable insights that go beyond this particular context since they addresses the issue of pronunciation acquisition in adult learners as a research area that is both under-represented and under-researched. With its detailed longitudinal focus, this empirical investigation not only fills an evident research gap but also represents an important point of departure for further research in a field that deserves due attention.

Secondly, the present results gear the scientific debate of the generalizability of the variables which influence the pronunciation abilities of adult learners more towards the individual as a unique combination of these factors seem to shape the success or failure of each and every learner. In this regard, no single individual factor could be identified as the major driving force in the language learning process. Rather, a dynamic combination of interrelated psycho(bio)logical factors like motivation to learn the language and to improve one's pronunciation, musicality, the lack of language anxiety as well as socio-cultural factors, most notably increased exposure to the target language (e.g. in the form of English-medium instruction) ostensibly have a positive effect on the development of the students' foreign accent. Yet predicting their impact at a specific point in time is highly problematic as their relative prominence is largely determined by the particular propensities of every individual learner, as the discussion of high achiever #61 and low achiever #29 has shown.

Taken together, these results lend support to the view that the scientific quest for a theoretical framework as all-encompassing as the Critical Period Hypothesis obscures reality and should eventually be abandoned in favour of a more individualised perspective of the learner in his/her socio-cultural, educational and – above all – psychological context. Clearly, future SLA research will have to grapple with the complex peculiarities of the pieces of a much larger puzzle.

References

Avello, P., Mora, J. C., & Perez-Vidal, C. (2012). Perception of FA by non-native listeners in a study abroad context. *Research in Language, 1*(1), 63–77. https://doi.org/10.2478/v10015-011-0050-9

Bernaus, M., Masgoret, A., Gardner, R., & Reyes, E. (2004). Motivation and attitudes towards learning language in multicultural classrooms. *International Journal of Multilingualism, 1*(2), 75–89. https://doi.org/10.1080/14790710408668180

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. F. Kroll & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.

Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition, 19*(04), 447–465.

Brantmeier, C., Schueller, J., Wilde, J. A., & Kinginger, C. (2007). Gender equity in foreign and second language learning. In S. Klein (Ed.), *Handbook for achieving gender equity through education* (pp. 305–334). Mahwah, NJ: Lawrence Erlbaum.

Campbell, C., & Ortiz, J. (1991). Helping students overcome foreign anxiety: A foreign language anxiety workshop. In E. Horwitz & D. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 153–168). Englewood Cliffs, NJ: Prentice Hall.

Christiner, M., & Reiterer, S. M. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology, 4*(874). https://doi.org/10.3389/fpsyg.2013.00874

De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition, 10*(1), 7–21. https://doi.org/10.1017/S1366728906002732

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*, 379–397. https://doi.org/10.2307/3588486

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition, 26*(2), 249–273. https://doi.org/10.1017/S0272263104262052

Díaz-Campos, M. (2006). The effect of style in second language phonology: An analysis of segmental acquisition in study abroad and regular-classroom students. In C. A. Klee & T. L. Face (Eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 26–39). Somerville, MA: Cascadilla Press.

Dogil, G., & Reiterer, S. M. (Eds.). (2009). *Language talent and brain activity. Trends in applied linguistics 1*. Berlin-New York: Mouton de Gruyter.

Dorfberger, S., Adi-Japha, E., & Karni, A. (2009). Sex differences in motor performance and motor learning in children and adolescents: An increasing male advantage in motor learning and consolidation phase gains. *Behavioural Brain Research, 198*, 165–171. https://doi.org/10.1016/j.bbr.2008.10.033

Dörnyei, Z. (1998). Motivation in second language learning. *Language Teaching, 31*(3), 117–135.

Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247–267). London: Continuum.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford, UK: Blackwell.

Elliot, A. R. (1995). Foreign language phonology: Field independence, attitude, and the success of formal instruction in Spanish pronunciation. *The Modern Language Journal, 79*(4), 530–542.

Flege, J. E., Munro, M., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125–3134.

Flege, J. E., Yeni-Komshian, G., & Liu, H. (1999). Age constraints on second language acquisition. *Journal of Memory & Language, 41*, 78–104.

Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition, 26*(2), 275–301. https://doi.org/10.1017/S0272263104262064

Gallardo del Puerto, F., Garcia Lecumberri, M., & Gomez Lacabex, E. (2009). Testing the effectiveness of content and language integrated learning in foreign language contexts: Assessment of English pronunciatio. In Y. Ruiz de Zarobe & R. M. Jimenez Catalan (Eds.), *Content and language integrated learning: Evidence from research in Europe* (pp. 215–234). Bristol, UK: Multilingual Matters.

Gardner, R., & Tremblay, P. F. (1994). On motivation, research agendas, and theoretical frameworks. *The Modern Language Journal, 78*(3), 359–368. https://doi.org/10.2307/330113

Gatbonton, E., Trofimovich, P., & Magid, M. (2005). Learners' ethnic group affiliation and L2 pronunciation accuracy: A sociolinguistic investigation. *TESOL Quarterly, 39*(3), 489–511.

Golestani, N., & Zatorre, R. (2009). Individual differences in the acquisition of second language phonology. *Brain & Language, 109*, 55–67. https://doi.org/10.1016/j.bandl.2008.01.005

Han, Z. (2004). To be a native speaker means not to be a non-native speaker. *Second Language Research, 20*, 166–187. https://doi.org/10.1191/0267658304sr235oa

Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism. Perspectives of change in psycholinguistics*. Clevedon, UK: Multilingual Matters.

Howard, M. (2005). Second language acquisition in a study abroad context: A comparative investigation of the effects of study abroad and foreign language instruction on the L2 learner's grammatical development. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 495–530). Berlin/New York: Mouton de Gruyter.

Isabelli, C. A., & Nishida, C. (2005). Development of the Spanish subjunctive in a nine month study-abroad setting. In D. Eddington (Ed.), *Selected proceedings of the 6th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 78–91). Somerville, MA: Cascadilla Press.

Jahandar, S., Hodabandehlou, M., Seyedi, G., & Mousavi Dolat Abadi, R. (2012). A gender-based approach to pronunciation accuracy of advanced EFL learners. *International Journal of Scientific and Engineering Research, 3*, 1–9.

Jilka, M. (2009). Assessment of phonetic ability. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity. Trends in Applied Linguistics* (Vol. 1, pp. 17–48). Berlin-New York: Mouton de Gruyter.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60–99.

Kenworthy, J. (1987). *Teaching English pronunciation*. Harlow, UK: Longman.

Khamkhien, A. (2010). Factors affecting language learning strategy reported usage by Thai and Vietnamese EFL learners. *Electronic Journal of Foreign Language Teaching, 7*(1), 66–85.

Khodadady, E., & Khajavy, G. (2013). Exploring the role of anxiety and motivation in foreign language achievement: A structural equation modeling approach. *Porta Linguarum, 20*, 269–286.

Kinginger, C. (2009). *Language learning and study abroad: A critical reading of research*. Basingstoke, UK: Palgrave/Macmillan.

Lafford, B., & Collentine, J. (2006). The effects of study abroad and classroom contexts on the acquisition of Spanish as a second language: From research to application. In B. Lafford & R. Salaberry (Eds.), *Spanish second language acquisition: From research findings to teaching applications* (pp. 103–126). Washington, DC: Georgetown University Press.

Lafford, B. A. (2006). The effects of study abroad vs. classroom contexts on Spanish SLA: Old assumptions, new insights and future research directions. In C. A. Klee & T. L. Face (Eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 26–39). Somerville, MA: Cascadilla Press.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*(2), 141–165.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Liu, M. (2006). Anxiety in Chinese EFL students at different proficiency levels. *System, 34*, 301–316. https://doi.org/10.1016/j.system.2006.04.004

Liu, M., & Huang, W. (2011). An exploration of foreign language anxiety and English motivation. *Education Research International, 1*(8), 1–8. https://doi.org/10.1155/2011/493167

Livingstone, S. (2002). *Young people and new media*. Thousand Oaks, CA: Sage.

Lord, G. (2006). Defining the indefinable: Study abroad and phonological memory abilities. In C. Klee & T. Face (Eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 40–46). Somerville, MA: Cascadilla Proceedings Project.

MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second language learning: Towards a theoretical clarification. *Language Learning, 39*, 251–275.

Mak, B. (2011). An exploration of speaking-in-class anxiety with Chinese ESL learners. *System, 39*, 202–214. https://doi.org/10.1016/j.system.2011.04.002

Marinova-Todd, S. F., Marshall, D. B., & Snow, C. (2000). Three misconceptions about age and L2 learning. *TESOL Quarterly, 34*(1), 9–31. https://doi.org/10.2307/3588095

Milovanov, R., Huotilainen, M., Valimaki, V., Esquef, P. A. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronujnciation skills in school-aged children: Neural and behavioural evidence. *Brain Research, 1194*, 81–89. https://doi.org/10.1016/j.brainres.2007.11.042

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition, 21*, 81–108.

Moyer, A. (2004). *Age, accent and experience in second language acquisition. An integrated approach to critical period inquiry*. Clevedon, UK: Multilingual Matters.

Moyer, A. (2007). Do language attitudes determine accent? A study of bilinguals in the US. *Journal of Multilingual and Multicultural Development, 28*, 1–17. https://doi.org/10.2167/jmmd514.0

Moyer, A. (2013). *Foreign accent. The phenomenon of non-native speech*. New York: Cambridge University Press.

Nakata, H. (2002). Correlations between musical and Japanese phonetic aptitudes by native speakers of English. *Reading Working Papers in Linguistics, 6*, 1–23.

Nakuma, C. (1998). A new theoretical account of 'fossilization': Implications for L2 attrition research. *IRAL, 36*(3), 247–256.

Nardo, D., & Reiterer, S. (2009). Musicality and phonetic language aptitude. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity. Trends in applied linguistics* (pp. 213–256). Berlin, Germany: Mouton De Gruyter.

Newport, E. (1990). Maturational constraints on language learning. *Cognitive Science, 14*, 11–28.

Oechslin, M. S., Meyer, M., & Jäncke, L. (2010). Absolute pitch – functional evidence of speech-relevant auditory acuity. *Cerebral Cortex, 20*, 447–455. https://doi.org/10.1093/cercor/bhp113

Ohata, K. (2005). Potential sources of anxiety for Japanese learners of English: Preliminary case interviews with five Japanese college students in the U.S. *TESL-EJ, 9*(3), 1–21.

Olivares-Cuhat, G. (2010). Relative importance of learning variables on L2 performance. *Linguistik online, 43*(3), 99–116. https://doi.org/10.13092/lo.43.415

Pastuszek-Lipińska, B. (2004). An overview of a research project and preliminary results of two experiments on perception and production of foreign language sounds by musicians and non-musicians. *TMH-QPSR, 46*(1), 61–74.

Patkowski, M. (1994). The critical age hypothesis and interlanguage phonology. In M. Yavas (Ed.), *First and second language phonology* (pp. 205–211). San Diego, CA: Singular Publishing Group.

Pickard, N. (1996). Out-of-class language learning strategies. *English Teaching Journal, 50*, 150–159.

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191–215. https://doi.org/10.1006/jpho.2001.0134

Preckel, F., Goetz, T., Pekrun, R., & Kleine, M. (2008). Gender differences in gifted and average-ability students. *The Gifted Child Quarterly, 52*, 146–159. https://doi.org/10.1177/0016986208315834

Price, M. L. (1991). The subjective experience of foreign language anxiety: Interviews with highly anxious students. In E. Horwitz & D. Young (Eds.), *Language learning anxiety: From theory and research to classroom implications* (pp. 101–108). Englewood Cliffs, NJ: Prentice Hall.

Purcell, E., & Suter, R. (1980). Predictors of pronunciation accuracy. *Language Learning, 30*, 271–287.

Rallo Fabra, L., & Juan Garau, M. (2010). Assessing FL pronunciation in a semi-immersion setting: The effects of CLIL instruction on Spanish-Catalan learners' perceived comprehensibility and accentedness. *Póznan Studies in Contemporary Linguistics, 43*, 96–108. https://doi.org/10.2478/psicl-2011-0008

Reid, J. (1995). *Learning styles in the ESL/EFL classroom*. Boston: Heinle & Heinle.

Reiterer, S. (2009). Brain and language talent: A synopsis. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity. trends in applied linguistics* (Vol. 1, pp. 155–191). Berlin-New York: Mouton de Gruyter.

Reiterer, S., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., et al. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology, 2*, 271. https://doi.org/10.3389/fpsyg.2011.00271

Richter, K. (2015a). EMI and pronunciation. A longitudinal case study of Austrian students at the University of Applied Sciences Vienna. Unpublished PhD thesis, University of Vienna.

Richter, K. (2015b). English Medium Instruction and Pronunciation. Outline of a longitudinal case study. In R. Wilkinson & M. Walsh (Eds.), *Integrating content and language in higher education: From theory to practice selected papers from the 2013 ICLHE conference* (pp. 275–290). Frankfurt, Germany: Peter Lang International Academic Publishers.

Richter, K. (2017). Researching tertiary EMI and pronunciation. A case study from Vienna. In J. Valcke & R. Wilkinson (Eds.), *Integrating content and language in higher education: Perspectives on professional practice* (pp. 117–134). Frankfurt, Germany: Peter Lang International Academic Publishers.

Ruiz de Zarobe, Y., & Jimenez Catalan, R. (Eds.). (2008). *Content and language integrated learning. Evidence from research in Europe*. Bristol, UK: Multilingual Matters.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York, NY: Newbury House/Harper & Row.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics, 20*, 213–223. https://doi.org/10.1017/S0267190500200135

Seither-Preisler, A., Parncutt, R., & Schneider, P. (2014). Size and synchronization of auditory cortex promotes musical, literacy and attentional skills in children. *The Journal of Neuroscience, 34*, 10937–10949. https://doi.org/10.1523/JNEUROSCI.5315-13.2014

Selinker, L., & Lakshmanan, U. (1992). Language transfer and fossilization: The multiple effects principle. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 197–216). Rowley, MA: Newbury House.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second language proficiency: Does music ability matter? *Psychological Science, 17*, 675–681. https://doi.org/10.1111/j.1467-9280.2006.01765.x

Smit, U. (2002). The interaction of motivation and achievement in advanced EFL pronunciation learners. *IRAL International Review of Applied Linguistics for Language Teaching, 40*(2), 89–115. https://doi.org/10.1515/iral.2002.009

Smit, U., & Dalton-Puffer, C. (2000). Motivational patterns in advanced EFL pronunciation learners. *IRAL International Review of Applied Linguistics for Language Teaching, 38*(3-4), 229–246. https://doi.org/10.1515/iral.2000.38.3-4.229

Stephenson Wilson, J. T. (2006). Anxiety in learning English as a foreign language: Its associations with student variables, with overall proficiency, and with performance on an oral test. Doctoral thesis. Universidad de Granada. Retrieved from http://hera.ugr.es/tesisugr/16235290.pdf

Tanaka, A., & Nakamura, K. (2004). Auditory memory and proficiency of second language speaking: A latent variable analysis approach. *Psychological Reports, 95*, 723–734. https://doi.org/10.2466/pr0.95.3.723-734

Tóth, Z. (2007). Predictors of foreign-language anxiety: Examining the relationship between anxiety and other individual learner variables. In J. Horváth & M. Nikolov (Eds.), *Empirical studies in English applied linguistics* (pp. 123–148). Pécs, Hungary: Lingua Franca Csopor.

Uskoski, O. (2011). *Playing video games: A waste of time… or not? Exploring the connection between playing video games and English grades.* Unpublished Pro Gradu thesis. University of Helsinki, Department of Modern Languages. Retrieved from https://helda.helsinki.fi/bitstream/handle/10138/35037/playingv.pdf?sequence=1

Woodrow, L. (2006). Anxiety and speaking English as a second language. *Regional Language Centre (RELC) Journal, 37*, 308–328. https://doi.org/10.1177/0033688206071315

Language Transfer vs. Language Talent? Individual Differences and Aptitude in L2 Phonology of Persian-Speaking Learners of English

Zhaleh Ghafoorian Maddah and Susanne M. Reiterer

Abstract Only little research regarding the phonology of Persian native speakers' English has been conducted. In the present study, we compared different individual cognitive factors which result in ESL Iranian English pronunciation, such as cognitive ability and short-term memory (working memory and Llama_D), language aptitude (MLAT III, IV and V; Llama_D) and working memory (Tewes U, Hamburg-Wechsler-Intelligenz-Test für Erwachsene Revision. HAWIE-R, Bern, 1994). These measures were correlated with English pronunciation and phonetic measurements (vowel length measurement) of Persian ESL learners. The sample comprised 30 Iranians aged 20–40 years (mean age 26.08) with L1 Farsi and an academic education. Their age of onset of learning (AOA) ranged from 2 to 16 years (mean age 11.03). Three learner groups were identified based on their language proficiency, and results confirmed previous findings about the contribution of cognitive factors (Rota G, Reiterer SM, Cognitive aspects of pronunciation talent. In: Dogil G, Reiterer S (eds) Language talent and brain activity. Mouton de Gruyter, Berlin, pp 67–96, 2009), language aptitude and multilingualism in L2 phonological processing. We observed significant correlations between English pronunciation scores and these three factors: schwa length pronunciation ($r = -0.8$), MLAT III ($r = 0.8$) and working memory ($r = 0.78$). Schwa length pronunciation also correlated highly with the number of languages ($r = -.74$) and the age of onset of acquisition ($r = .41$). Our cross-linguistic results suggest that phonological native-like L2 achievement in ESL adult learners is possible, as individuals with higher L2 aptitude and working memory capacity can overcome the transfer of L1 phonological categories in L2 processing.

Z. G. Maddah (✉) · S. M. Reiterer
Department of Linguistics and Centre for Teacher Education, Unit for Language Learning and Teaching Research, University of Vienna, Vienna, Wien, Austria
e-mail: a0448481@unet.univie.ac.at

S. M. Reiterer (ed.), *Exploring Language Aptitude: Views from Psychology, the Language Sciences, and Cognitive Neuroscience*, English Language Education 16, https://doi.org/10.1007/978-3-319-91917-1_19

1 Introduction

Second language learning has been the focus of many pedagogical and linguistic investigations so far. It has been shown that productive skills, speaking and writing, are more difficult for the second language learners than receptive skills, listening and reading. Native-like proficiency, e.g. in pronunciation, is reached only by approximately 5–15% of late language learners (Reiterer et al., 2011; Selinker, 1972). Accordingly, mastering the phonological system of a foreign language can be said to be unlikely at a native-speaker level in adulthood (Cabrelli Amaro & Rothman, 2010, p. 277). However, some adult L2 learners can overcome their age-related constraints and achieve a near-native L2 accent due to their individual traits (Erard, 2012).

There are great individual differences observed in the perception and production of ESL learners, which is why this paper focuses on individual differences in second language (hereafter L2) phonology. Specifically, speech production will be investigated, whereupon the following components will be examined: vowel production (vowel length), speech imitation of an unknown language (L0), language aptitude and cognitive ability. The main goal of this study is to find out if there is any relationship between language aptitude tests and the proficiency in English pronunciation in terms of English native-likeness and (schwa duration) as a phonetic marker.

The Persian or Farsi language has its own sound system with some differences to the English language. These differences are reflected in the speech of Persian speakers of English and result in phonological errors when speaking English (L1 transfer). It is shown that L1 transfer is not a general phenomenon applicable to every speaker. Affected are the medium range (70% of a population, one standard deviation above and below the mean) and the lowest quantiles, however not the highest quantiles (5–15%) of speakers with high language aptitude.

In this study, we investigated Persian native speakers with no specific immersion into the English language. As Farsi has a different phonetic inventory than English, we took this L1 population to test, assuming that the subjects are affected individually differently by L1 transfer in their English pronunciation. In this regard, it would be useful to have a brief look at different transfer models in second language acquisition.

2 Transfer Models in Second Language Acquisition

Transfer in learning was investigated by psychologists roughly 85 years before the introduction of Contrastive Analysis (CA). Transfer in SLA was ignored during the 1970s due to shortcomings of CA, but from the 1980s onwards it gained popularity again (Major, 2008, p. 65). As Ausubel (1963, p. 28) mentions, transfer is involved in all kinds of learning, provided that there are some "relevant aspects" in experience and that it is "organically relatable", something which was earlier referred to as having "meaningful similarity" by Osgood (1946). Put in simple terms, transfer occurs due to "connections between old and new information" (Neuner, 2002.).

Transfer can happen at different levels of language such as phonology (sound systems), morphology (word structure), syntax (sentence structure) and lexical semantics (Thomason, 2001). Also, Goad and White (2006) consider the transfer of the "prosodic system" of L1 as one of the main factors in the acquisition of native-like proficiency in L2.

Theories on language transfer are described with 3 approaches: firstly, the Full Transfer or Full Access hypothesis which asserts that because of the L1 influence, learning a foreign language after a certain period is almost impossible; secondly, theories which state that the L1 transfer occurs partially – in this view, some subfields of language are more prone to transfer than others; thirdly, theories focusing on individual differences in language transfer regarding aptitude and environmental factors. Furthermore, as Wardhaugh (1970) mentions, language transfer can be positive or negative: positive transfer occurs when L1 and L2 linguistic items are similar and as a result facilitate learning; negative transfer hinders L2 acquisition because of a mismatch in L1 and L2.

3 Language Aptitude, Some Categorizations and Tests

Language aptitude is a concept which is related to the broader concept of human cognitive abilities, covering a variety of cognitively-based learner differences. Such differences affect any kind of learning in individuals including L2 learning.

Language learners differ in their ability to learn a second or foreign language. According to Wen (2011, p. 233), it is shown that some learners are able to learn a foreign or second language more easily and quickly that other people. Such ability has been mentioned also synonymously with 'language aptitude', a special 'propensity' or 'talent' for learning an L2 and a 'flair' or 'knack' for languages (Dörnyei, 2005). John B. Carroll refers to language aptitude as "simply an ability or "knack" for learning foreign languages (Dörnyei, 2005).

Language scholars proposed different models of foreign language aptitude on the basis of their categorization of cognitive abilities. Robinson (2002) considers cognitive abilities as hierarchical abilities in that each order contributes to the abilities in the next order. For example, first order abilities refer to abilities that are measured by psychological tests such as "working memory capacity" and "analogical reasoning". On the other hand, second order abilities are abilities which result from a special combination of first order abilities (e.g. "broad intelligence" and "fluid speediness").

Carroll's model of foreign language aptitude (Carroll, 1981; Carroll & Sapon, 1959) categorizes the skills for L2 learning accordingly: "phonemic coding ability (memory of sounds and their combinations), associative memory (the ability to remember new words), inductive language learning ability (the ability to find patterns in words and sentences), and grammatical sensitivity (the ability to understand sentence structure of unknown languages)". Carroll and Sapon (1959) also mention rote memorization ability as one of the subcomponents of their classical model of foreign language aptitude. In this regard, Reiterer et al. (2011) point out that a person can either have a “talent for accent” (Obler & Fein, 1988; Skehan, 2011) which can be related to Carroll’s “phonemic coding ability” or a “talent for grammar” (Nauchi & Sakai, 2009) corresponding to Carroll’s “inductive language learning ability”. Another model of language aptitude is proposed by Skehan (2002) who categorizes the subcomponents of foreign language aptitude in regard to cognitive factors such as noticing, patterning, controlling and lexicalizing. Additionally, Robinson (2002) defines cognitive resources as three kinds of memory (working memory, short-term memory and long-term memory) plus attention and basic processing speed. His model (2002; 2007) is called Aptitude Complex Hypothesis (ACH) which is based on aptitude complexes in instructional contexts with the aim of maximizing pedagogical performance of L2 learners. This model considers mainly L2 processing of L2 learners and their focus of attention and intention with L2 tasks.

The history of language aptitude tests goes back to the 1920s and 1930s in the United States when failing foreign language courses at school was common because the school program dedicated little time to foreign language study. As a result, the education system invested in the design of ‘prognosis tests’ in order to detect potential ‘causalities’ (Spolsky, 1995, cited in Dörnyei, 2005, p. 34). Between 1925 and 1930 three tests were designed which did not have specific ‘theoretical foundation’ but were based on two shared approaches for the measurement of language aptitude. Such tests in Spolsky’s view are categorized as analytical and synthetic with the former testing special cognitive abilities that are carried out in the students’ first language and the latter containing ‘mini tasks’ that are centered in learning an artificial foreign language or a rare existing L2. After 30 years between the 1950s and 1960s, a period which, in Rees’ (2000) words, is referred to as the “golden period’ of scientific language aptitude testing”, two ‘systematic tests’ were developed, one by John Carroll and Stanley Sapon, and one by Paul Pimsleur (1966). The test designed by John Carroll and Stanley Sapon is *The Modern Language Aptitude Test* (MLAT). This test was conducted on roughly 5000 participants at Harvard University from 1953 to 1958 and aimed to predict the accomplishment in foreign languages (Carroll & Sapon, 1959). *The Modern Language Aptitude Test* (MLAT) comprises five sub-tests, namely the MLAT I, II, III, IV and V. Each test aims at measuring different subcomponents of language aptitude according to Carroll’s model of foreign language aptitude (Skehan, 1998). These are phonemic coding ability, associative memory, inductive language learning ability and grammatical sensitivity. For the present study, we administered the three last parts, namely MLAT III, IV and V.

4 Research Questions and Hypotheses

The present study started with two central research questions:

(a) can adult L2 learners produce L2 phonological features in a native-likemanner?

(b) to what extent do language aptitude, and cognitive and mimicking abilitycontribute to the establishment of native-like phonetic attainment?

To answer these questions, we formulated six hypotheses, considering that factors such as phonological transfer, age, gender, language aptitude, schwa pronunciation length, imitation ability and memory would have an effect on native like L2 attainment. The hypotheses were formulated as follows:

H1: ESL learners transfer some parts of their mother tongue's phonological. categories into their English pronunciation.
H2: An earlier age of onset (AOA) of language influences L2 learning in a positive way.
H3: The subjects with better English pronunciation articulate the schwa sound shorter.
H4: The ability to imitate an unknown language is related to L2 learning aptitude.
H5. Females are better language learners than males.
H6: Higher L2 aptitude and a better working memory are interrelated.

5 Methodology

This part of the current research deals with subjects selection criteria and specific tests for measuring English pronunciation ability, imitation ability, working memory ability and language aptitude of the participants.

5.1 Tests

To assess the language aptitude of the subjects in this study, different tests were administered to all participants, namely four language aptitude tests (the MLAT III, IV and V, and the Llama_D test) and one cognitive ability test [Working Memory hereafter (WM) (Tewes, 1994)]. Because of the lack of schwa in the Persian sound system, we decided to focus on schwa pronunciation in our data analysis. For testing the subjects' English pronunciation, two short stories with schwa-containing words were used.

Oral English data were used for two purposes: first, for measuring English pronunciation in terms of the degree of native-likeness and second, for the purpose of phonetic analyses (schwa duration). Finally, an additional imitation ability task (Reiterer et al., 2011; Reiterer, Hu, Sumathi, & Singh, 2013) was added to the study to evaluate the subjects' ability of imitating an unknown language (L0), namely Hindi, in order to find out to what extent the participants can reproduce an unknown phonological stimulus. For this purpose, Hindi sentences with different syllables were embedded into the task procedure of the recordings.

5.2 Subjects

We selected a representative sample of Iranian ESL learners and recruited them for voice recordings. The subjects were English learners who were voluntarily taking ESL courses in private language institutes in Iran, with academic backgrounds ranging from B.A. to PhD. Some of the participants were learning English in IELTS preparation courses in order to continue education abroad or to emigrate to Canada. The subjects' age was in the age range of 20–40. There were 30 participants in total (N = 30) with 18 females and 12 males. The average age of females was 25.83 years (SD = ±4.719, Min 20, Max 34) and the average age of males was 28.25 years (SD = ±3.957, Min 24, Max 39). Females were on average 2.42 years younger than males.

Additionally, the participants (a) had completed their study at high school, (b) were studying or have finished their study at the university in B.A, Master or PhD levels, (c) had no prior contact or immersion in Hindi language, (d) were all native speakers of Persian raised by monolingual parents, and (e) had no significant exposure to English before the age 12 (this is the age, 2nd grade middle school, in which students start to learn English at school). On average, they reported to have studied English for approximately 16 years (range 5–28 years). They all reported to use English mainly for educational purposes and not to communicate with native speakers of English. Almost all of the participants were living in the North East of Iran with no particular immersion into the English language.

5.3 Questionnaire

The questionnaire was designed to get some information regarding the subjects' personal and linguistic background data, such as date and place of birth, age of onset of language learning, exposure to any other foreign language than English, number of languages spoken, time spent in English speaking countries, number of dialects, type of exposure to the English language and education level.

5.4 Phonological Data

In the phonological measurements we focused on one of the acoustic correlates of vowel quality, namely the duration or vowel length. Other acoustic properties which can also be examined are e.g., voice onset times (VOT), fundamental frequency (F0), first formant (F1) and second formant (F2) (Rosch, 1975), ratios or distances between formants.

Two sets of recordings were planned for measuring pronunciation and imitation abilities of the subjects. Firstly, read aloud tasks, where the informants had to read out two fables namely "The North wind and the Sun" (Aesop fable) and "The Lightning" (short story by Marc Twain) in their best accent possible (American or British English).

They were given 10 min to prepare for the reading. Individual recordings of the first short story took about 1 min and the second story 2 min. Depending on the participants' reading pace, the time was slightly above or below this threshold. After finishing the English pronunciation task, in the second set of recordings, every subject was asked to listen to 4 Hindi sentences, read out by a native Hindi speaker, 3 times each, and imitate the sentence afterwards as close to the original sentence as possible. In this part of the recording, there were just phonological stimuli and no lexical stimuli contrary to the English pronunciation task where subjects had to read from a printed text. The recordings (English pronunciation) were used for two purposes: first for ratings by native English speakers and second for phonetic analyses. The recorded Hindi data were used to measure the imitation ability of an unknown language.

For our analysis of the English recordings, the recorded data were digitalized. We measured the schwa duration of initial and inside (mid position) schwa sounds in the function and lexical words of the English texts, which amounted to 28 words for each subject. We measured one schwa per word i.e. 28 schwas for each participant. For the analysis, the computer program 'Adobe Audition Professional' was used. The version was compatible with a Windows operating system. Digital sound files of the recordings in MP3 format were used. In total, these were 60 sound files (30 subjects, each reading two short stories). A headphone was employed to reach the maximum concentration on the task as this analysis required careful listening. Every file was played from the beginning up to each schwa containing word. Then the waveform of the word containing the schwa was maximized to detect the schwa visually. In order to do this, each schwa containing word was played over and over again to identify the exact position of schwa boundaries in the waveform. Next, the duration of a schwa token was computed as the time between its starting and end point measured in milliseconds (ms) as determined by the software.

For the rating of the subjects' performance in English pronunciation four adult native English listeners rated the English recorded data for the overall impression of native-likeness using a 10-point Likert scale. Hindi sentences were also rated by 4 native Hindi adult listeners for the subjects' performance on the Hindi imitation task using a 10-point Likert scale. The four raters per language (English and Hindi) were given an assessment sheet with the participants' initials (in each rater's sheet the order of the subjects was different). Each rater was asked to listen to the recordings and note down their score using a scale from 0 (the poorest performance) to 10 (a native-like performance). Indication of decimals, e.g. 9.5, was also possible. The average of the scores of English and Hindi were calculated for each subject. The participants were split into three parts according to their performance in the English pronunciation task: i.e. good pronunciation (n = 10), average pronunciation (n = 10) and poor pronunciation (n = 10).

6 Results

The following subsections deal with the results of the phonological data and other test scores and look into their relationships with one other.

6.1 Phonological Data

After the completion of the ratings of English pronunciation tasks, scores were added and means calculated for each participant. The scores ranged from 2.50 to 9.50 points. Hindi scores ranged from 2 to 8 points. As is shown in the correlation, there is a *significant positive* correlation between the scores in the Hindi imitation task and English pronunciation ability ***r* = .36,** p (two-tailed) < .05 which confirms our hypothesis that better English pronunciation is related to a better L0 (here Hindi) imitation ability. The inter-rater concordance / reliability was *r* = .67 (p = .000) for both English texts (4 raters) and *r* = .79 (p = .000) between the 4 Hindi native raters.

This part of the data addresses the two central research question of this study namely: (a) can adult L2 learners produce L2 phonological features in a native-like manner? and (b) to what extent do language aptitude, and cognitive and mimicking ability contribute to the establishment of native-like phonetic attainment? In this regard, English pronunciation and Hindi imitation scores were correlated with each other and also with other factors. We considered scores between 8 and 10 to be in the native speaker range. The following histograms show the distribution of the scores of the English and Hindi tasks (Figs. 1 and 2).

6.2 Comparing Males' and Females' Performances on Hindi and English Tasks

The result of Hindi imitation and English pronunciation abilities shows that females scored, on average, higher than males. The average score of females in Hindi imitation task was 6.3 with SD = ±1.20 (Min 2.79, Max 7.78). Mean of males' scores was 4.8 with SD = ±1.66, (Min 1.82, Max 7.29). Regarding English pronunciation ability, the mean score of females was 5.77 (SD = ±1.85, Min 3.5, Max 9.31). The average of males' scores was 4.43 (SD = ±1.83, Min 2.5, Max 7.69). The comparison shows that males scored roughly the same in English (M 4.43) and Hindi (M 4.80) tasks with 0.37 points difference. Their maximum scores in both task was also similar with 0.30 points difference. Females scored on average 1.49 points higher in Hindi and 1.34 points higher in English scores than males. Also, the Kolmogorov-Smirnov test shows that the English and Hindi scores are normally distributed.

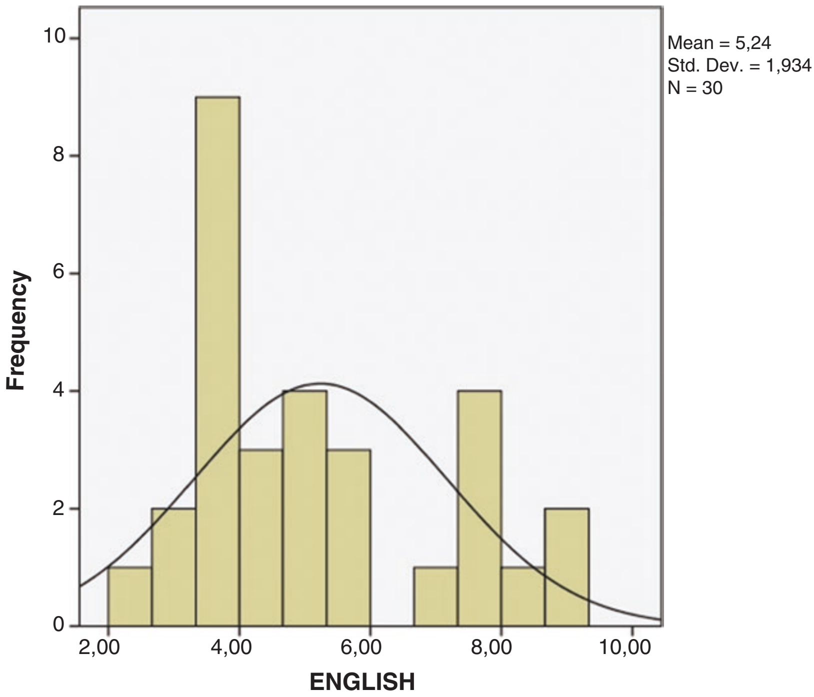


Fig. 1 Histogram-Distribution of the subjects with respect to English pronunciation score

6.3 Phonetic Analysis: Measuring Vowel Length Duration

Vowel length analysis revealed significant differences in the performances of different ability groups, as mentioned before, subjects were grouped according to their scores by English native speakers in the English reading task: three groups were identified: low, medium and high ability (N = 10 each). It was found that less proficient subjects tended to pronounce the schwa longer and there was a clear trend in the increase of schwa pronunciation length with a decrease in English pronunciation scores (Fig. 3). As expected, there was a significant difference in schwa length pronunciation between the high and low ability groups. As the bar chart (Fig. 4) shows, the low ability group's average of schwa duration was 110 milliseconds and the high ability group pronounced the schwa around 75 milliseconds on average which shows a significant difference of 35 milliseconds.

Schwa duration length in the subjects' performances was in the range of mid 40 ms–250 ms, with the talented group being very near to the native speakers' duration. In case that one word contained more than one schwa this was specified with the extension "2″ (see examples below (*ashamed-2*)). The Figs. 5 and 6 illustrate the oscillogram of schwa duration pronunciation of a high ability non-native speaker (Fig. 5) and a native speaker (Fig. 6) likewise.

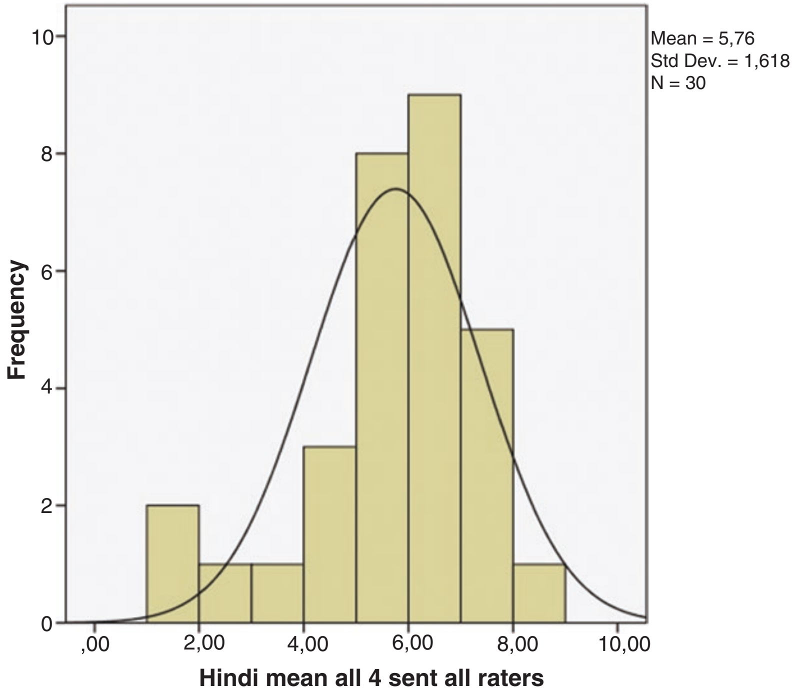


Fig. 2 Histogram-Distribution of the subjects with respect to Hindi imitation score

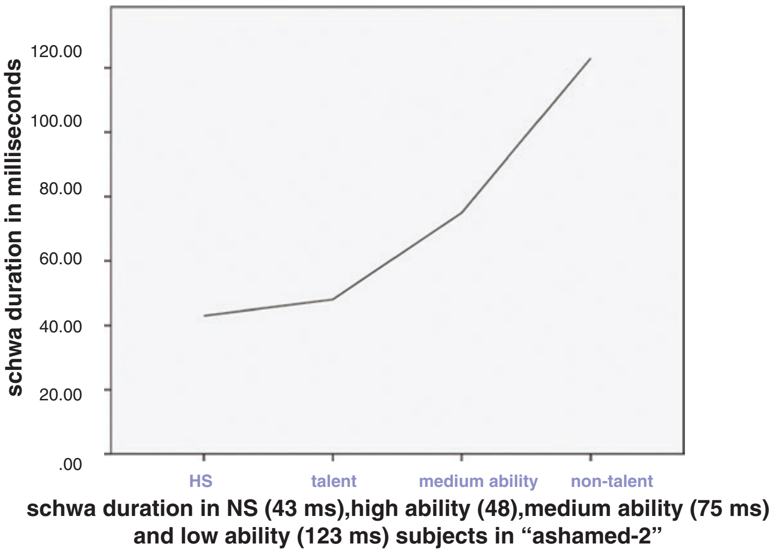


Fig. 3 Schwa length pronunciation in *ashamed-2* in different ability groups (milliseconds values)

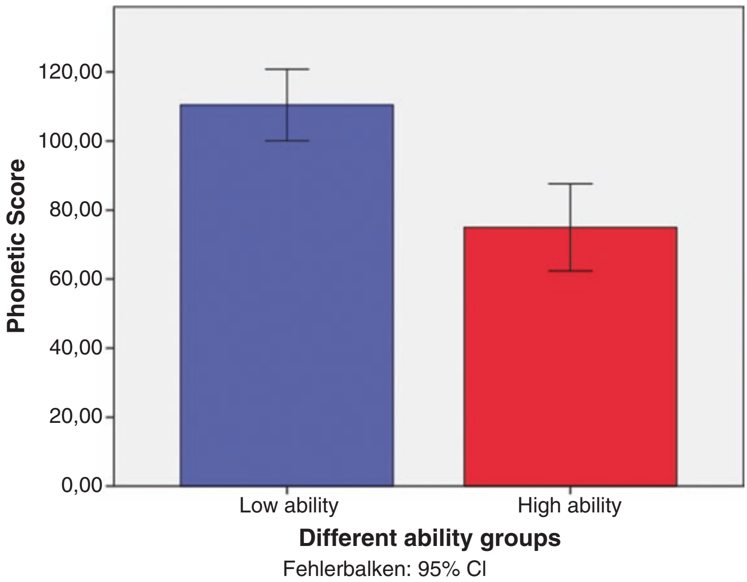


Fig. 4 Schwa length pronunciation in extreme ability groups

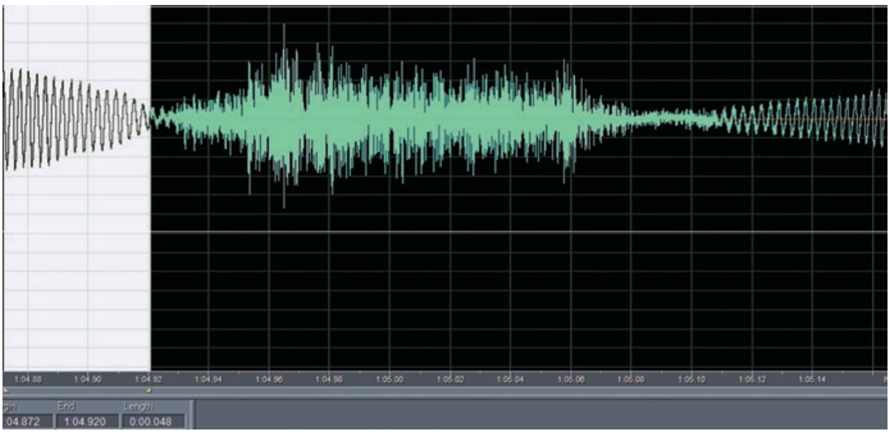


Fig. 5 Schwa duration of a high ability non-native speaker in the word *ashamed – 2*

A t-test analysis illustrates that the mean schwa duration (phonetic score) is: 75 milliseconds (SD = ±17.66, high ability group), and 110 ms (SD = ± 14.5, low ability group). The difference between high and low ability groups' phonetic scores is 35.5 milliseconds and is highly significant (p = .000).

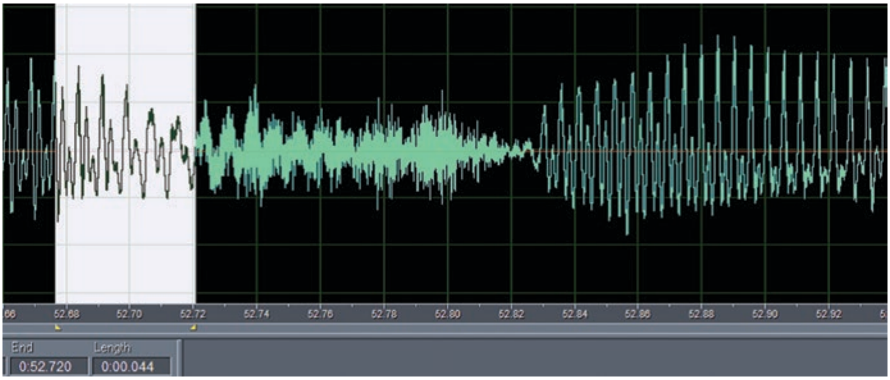


Fig. 6 Schwa duration of English native speaker in *ashamed – 2*

6.4 Correlation of Schwa Length Pronunciation with Other Factors

The correlation analysis of schwa length pronunciation with other factors contributing to ultimate L2 attainment revealed significant results. The subjects with more native-like English pronunciation articulated the schwa sounds with shorter duration. Other variables such as the Llama_D test score, Language aptitude tests results, number of foreign languages spoken and working memory also showed significant correlations.

The scatter plot in Fig. 7 depicts one of the main findings of this study. The graph below shows the strong negative relationship between the duration of the schwa and the native-like pronunciation score of the participants by the English native raters: the higher the scores, the shorter the schwa. As expected, there is a negative, linear relationship ($r = -.8$) between the English pronunciation score and schwa duration (phonetic score). Thus, the shorter the schwa sound is pronounced, the higher the score in English pronunciation will be.

Assuming that the ability to imitate an unknown language is related to L2 learning aptitude, it was expected that participants who pronounced schwa shorter got better scores in their Hindi pronunciation task. And indeed we found a pronounced negative correlation between the schwa length pronunciation and Hindi imitation scores at trend level ($r = -.35$, $p = .057$) in all ability groups. This result confirms that better imitation talent is positively related to more native-like performance, here in terms of a shorter pronunciation of schwa.

The correlation of schwa length duration with the age of onset of the language was significant. There was a positive correlation between age of onset of language learning and the length of schwa duration ($r = .41$, p (two-tailed) $< .05$). This result suggests that the later ESL learners start to learn a foreign language, the less they can achieve native-like pronunciation. In other words, the longer the subjects pronounced the schwa, the older they were when they started learning English. This finding is in line with Ioup (2008) who mentions that “phonological accents in a

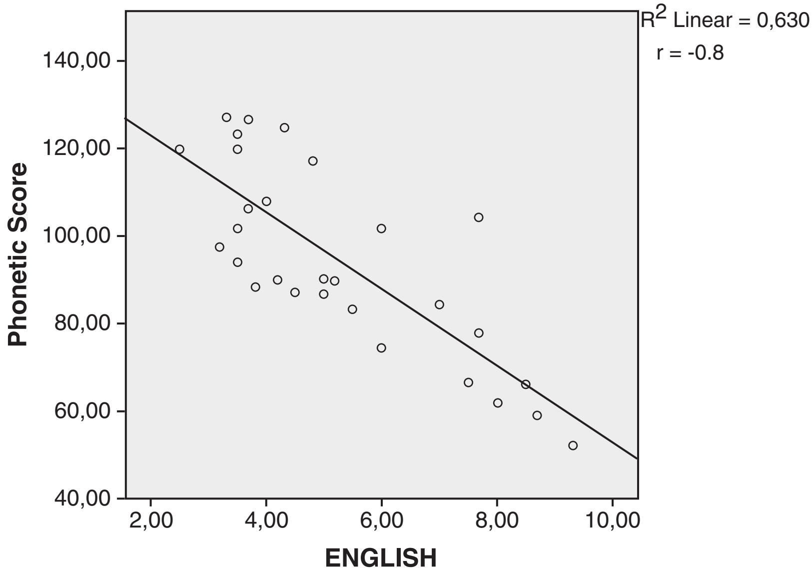


Fig. 7 Correlation between English pronunciation score and phonetic score

second language (L2), more than other linguistic skills, would more exhibit age effects, because accent was the only part of language that was physical and demanded neuromuscular programming".

Another significant result of the present study was the relationship between the number of foreign languages spoken and schwa length pronunciation: negative correlation between the number of languages spoken and schwa length duration, $r = -.74$, p (two-tailed) < .01. Thus, the data suggest that the more languages a person speaks, the shorter they tend to pronounce schwa and can imitate a foreign language sound more efficiently.

The second central research question of this study was "to what extent do language aptitude, and cognitive and mimicking ability contribute to the establishment of native-like phonetic attainment?" In the scatter plot below, a significant, negative correlation ($r = -0.6$) can be observed between working memory scores and schwa length pronunciation. As it is illustrated, participants who pronounced schwa shorter achieved higher WM scores and vice versa (Fig. 8).

6.5 Correlation Between English Pronunciation Score and English Aptitude Tests

To see in how far English aptitude tests can predict the language learning ability of ESL leaners, we looked into the results of these tests and the subjects' performance in different tasks. English aptitude test results of the participants confirmed that there is

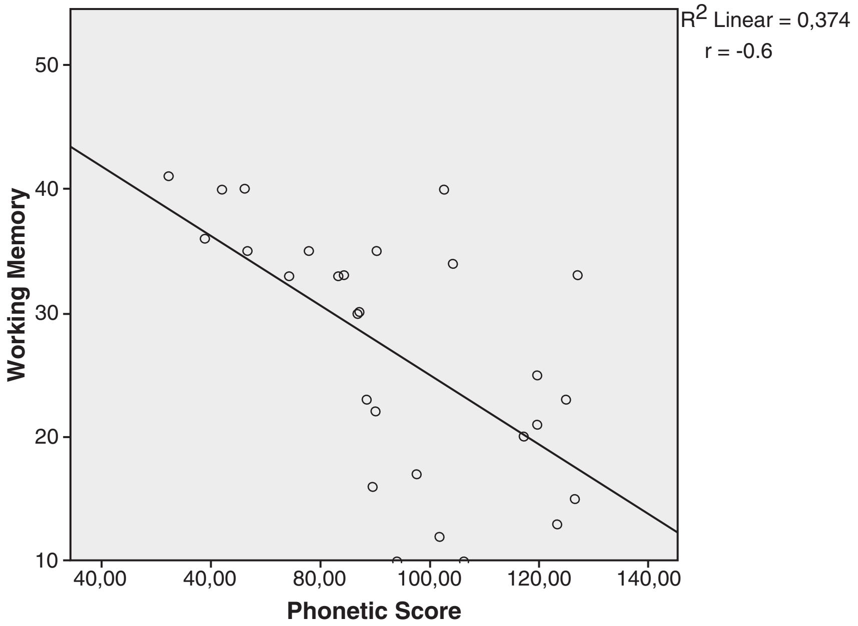


Fig. 8 Correlation between Phonetic Score (SCHWA-vowel duration) and working memory

a significant association between the scores of language aptitude tests and a person's native-likeness in ESL. As mentioned above, subjects were tested in three MLAT battery sub-tests, namely the MLAT III, IV and V. The data shows that the most significant correlation can be observed with MLAT III (phonetic coding) test scores followed by MLAT IV and V scores. As illustrated in Fig. 9 below, there is a significant, positive correlation ($r = .8$) between MLAT III and English pronunciation scores.

MLAT IV also correlated highly with the English pronunciation and scores showed the sig. correlation between English pronunciation and MLAT IV test scores (Pearson $r = 0.73$) in all ability groups. Likewise, the MLAT V score correlated highly with English pronunciation (Pearson r = .63).

The results for the relationship between English pronunciation scores and the Llama_D test scores, which was the last L2 aptitude test used in this study, show a positive correlation between the scores of English pronunciation and the Llama_D test scores (Pearson $r = 0.65$).

6.6 Correlation Between Working Memory and Other Factors

As mentioned before, the only cognitive ability test which was implemented in the present study was a working memory test. In this section, we examined the correlation of the WM test with other variables such as English pronunciation score, schwa

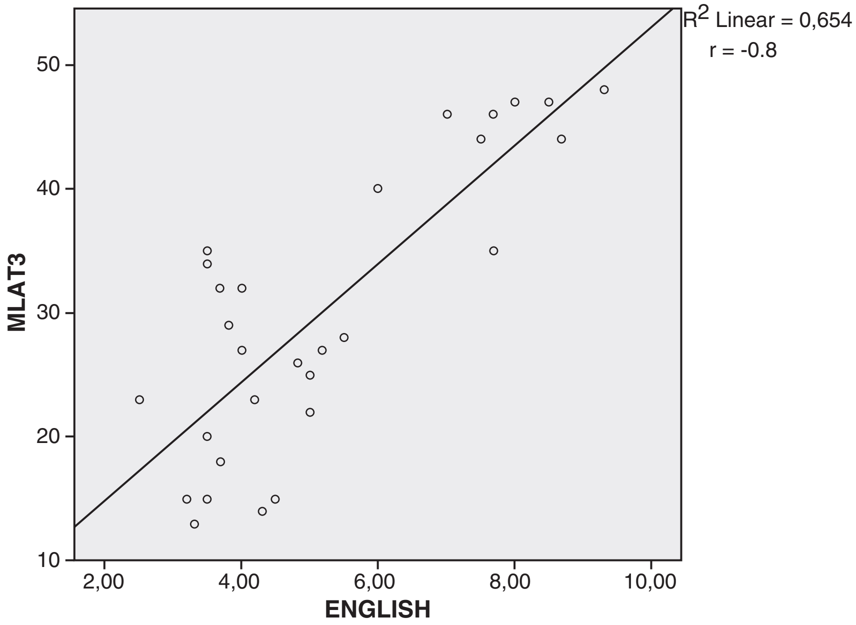


Fig. 9 Correlation between English pronunciation score and MLAT III

length pronunciation and language aptitude tests. We hypothesized that WM capacity is an effective factor in native like attainment (H6). The results of WM tests revealed significant correlations at $p < .01$ level with all the variables considered in this study in order to achieve a native-like level.

Figure 10 illustrates the relationship between the scores of English pronunciation and the scores in the WM test. As the graph depicts, there is a strong correlation between English pronunciation scores and the performance in the WM test, $r = 0.78$ which shows that ESL speakers with higher WM capacity also performed better in English pronunciation.

There are two significant results with respect to language aptitude and the cognitive tests of this study. Firstly, it was shown that females could outperform males in aptitude and cognitive tests (Fig. 11). Secondly, the results of all tests correlate significantly with each other. For example, the correlation between MLAT III and WM is $r = .628$, and the correlation between MLAT III and MLAT IV is $r = .661$. Accordingly, the data suggest that the performance in one test could be a reliable predictor for the result of another test. Hence, a person with a higher WM is expected to perform better in MLAT III and a good result in MLAT III suggests a better result in MLAT IV.

Extreme group comparison shows that there is a significant difference between the performances of both groups in all tests (Fig. 12). Due to the comparison with other tests, the scores of the Llama_D test multiplied by 100 as the test results of the Llama_D are presented in percentages.

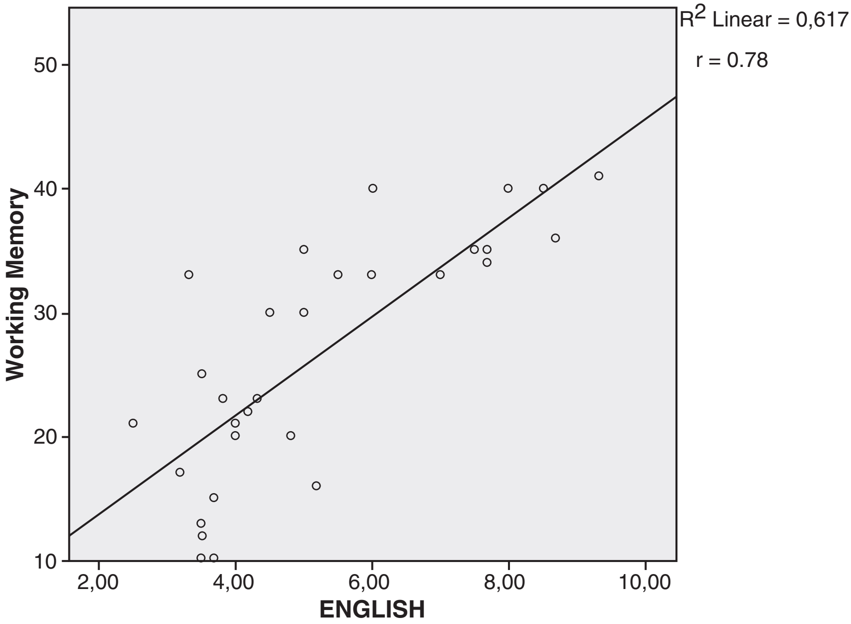


Fig. 10 Correlation between English pronunciation score and working memory

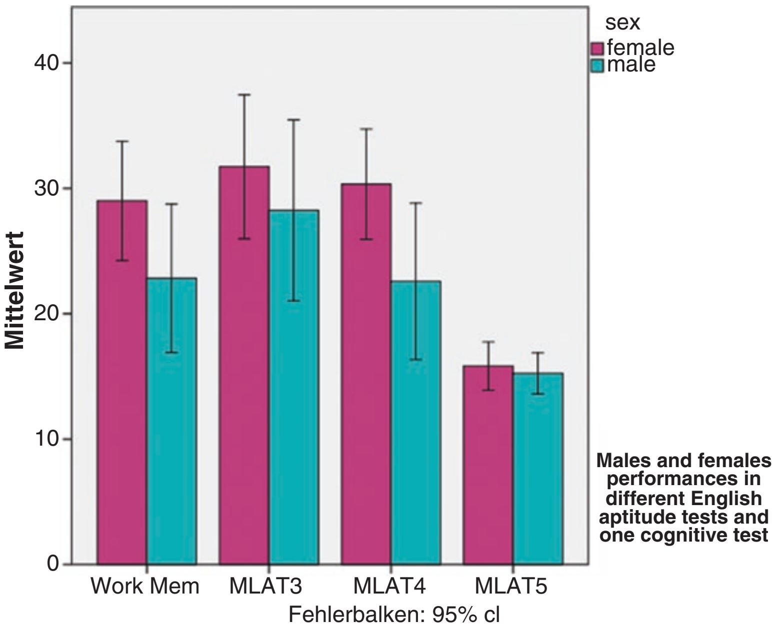


Fig. 11 Males' and females' performances in MLAT 3, MLAT 4, MLAT 5 and working memory

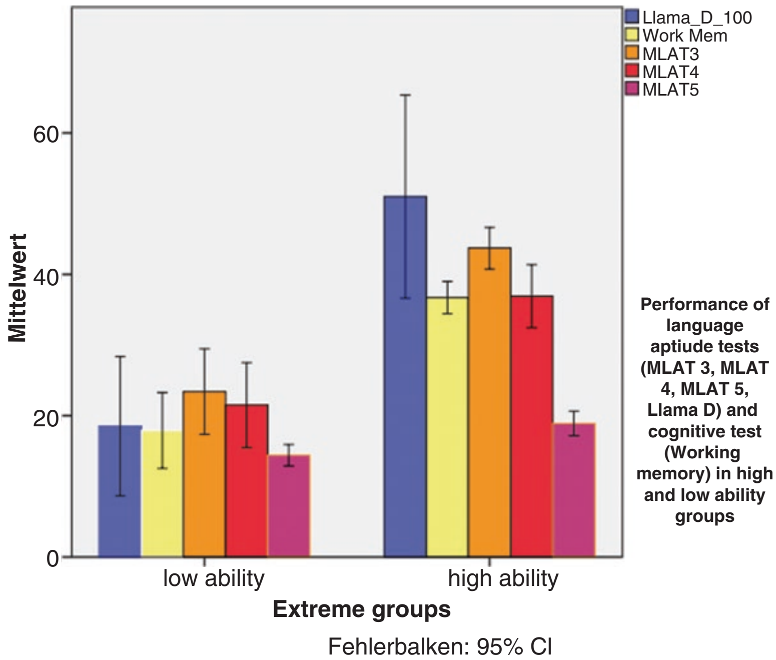


Fig. 12 Group differences of language aptitude tests (MLAT 3, MLAT 4, MLAT 5, Llama_D) and cognitive test (working memory) in extreme groups

The bar charts (Fig. 12) compare the result of aptitude and cognitive tests in different ability groups. As expected, the result of the talented group is significantly higher than the medium and low ability groups' results in all tests, especially in the MLAT III test. However, in the MLAT V test, different ability groups did not perform so differently. Even in this case, the low ability group was slightly better than the medium ability group. In addition, the result of MLAT III and IV show no significant difference between these two groups. Working memory test results (Fig. 13) showed a clear-cut difference in performance between all ability groups including medium and low abilities which was not observed in other test results.

7 Discussion

Phonological aptitude presents an interesting case study for L2 research because it is the area in which L2 learners have most difficulties. Native-like accent in L2 has been the ultimate goal in foreign and second language learning. Moreover, mastery in L2 phonology and L2 native-like accent is an exception rather than the norm and is achieved by only a handful of L2 learners. The present study examined how L2 phonological proficiency is impacted by aptitude and cognitive abilities of ESL

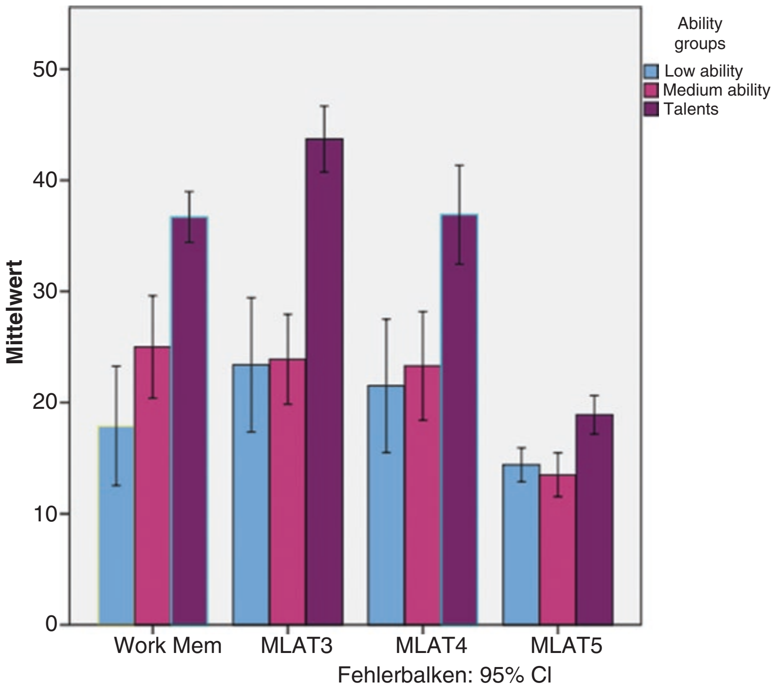


Fig. 13 Different ability groups' performances in MLAT III, MLAT IV, MLAT V and working memory

learners by investigating the relationship between English pronunciation and foreign language imitation scores. Additionally, linguistic-phonetic talent in Persian ESL learners was measured for the first time.

Two central research questions of this study were "can adult L2 learners produce L2 phonological features in a native-like manner?" and "to what extent do language aptitude, cognitive and mimicking abilities contribute to the establishment of native-like phonetic attainment?" We addressed these questions by administering a combination of various language aptitude tests, and subjects were tested on their English accent and imitation abilities. Since the L0 imitation task did not involve any prior experience with Hindi as a foreign language, we regard this task as a non-word repetition task (Munson, Edwards, & Beckman, 2005). The result of our study showed significant correlations between native-like English pronunciation ability, language aptitude, cognitive ability, L0 imitation, L0 recognition (Llama_D), vowel duration as phonetic marker, the age of onset of language learning and the number of foreign languages spoken. Particularly, it was observed that ultimate attainment in L2 mostly results from individual and cognitive differences, such as language aptitude and memory. Moreover, female participants performed better than males in English pronunciation, language aptitude tests and imitation tasks.

In our analysis, we found that the shorter pronunciation of the vowel schwa is an indication of native-likeness in English pronunciation in regard to phonetic talent

Table 1 Correlation between Hindi imitation task and English pronunciation score

Correlations		English	Hindi
ENGLISH	Pearson Correlation	1	,365*
	Sig. (2-tailed)		,048
	N	30	30
HINDI	Pearson Correlation	,365*	1
	Sig. (2-tailed)	,048	
	N	30	30

[a]Correlation is significant at the 0.05 level (2-tailed)

(H3). As it was presented before, this is a particular feature which is missing in the sound system of Persian. Hence, schwa pronunciation can be a reliable phonetic indicator of the mimicking ability of Persian ESL learners. The phonetic analysis showed that subjects who pronounced the schwa shorter achieved better scores in English pronunciation from English native speakers. The scoring of the native Hindi raters had to do with their perceived impression of native-likeness and not with their attention to schwa duration. Schwa duration measurements were carried out independently which was found to have a strong correlation with English pronunciation scores ($r = -0.8$). Based on previous research (Scovel, 1998) and the CPH (Penfield & Roberts, 1959; Lenneberg, 1967), it was hypothesized that native Persian speakers of English, especially late age of onset subjects, would transfer their L1 phonology system to the L2 English due to anatomical and biological constraints. Also, the usual L2 phonological theories would predict the transfer of the L1 system (leading to foreign accent). However, talented Iranian L2 learners who have successfully achieved near-native phonological proficiency without specific immersion into English showed that transfer theories do not equally apply to all individuals. Speech production data and results of language aptitude tests showed a significant impact of language aptitude in L2 learning ability.

Our study mainly focused on phonological aspect of L2 learning and native-like pronunciation attainment. To measure this construct of L2 learning, we tested subjects on English native-like accent (rating by native English raters), mimicking ability (Hindi imitation) and the MLAT III which is the only subpart of the MLAT battery measuring phonetic coding ability. Two most salient results regarding phonetic aptitude were the relationship between English proficiency scores with both the schwa length pronunciation and MLAT III. It was shown that:

(a) as the duration of schwa increases, scores of English pronunciation decrease ($r = -.8$),
(b) as the score on the MLAT III increases, the score on English pronunciation increases ($r = .8$).

This observation suggests that schwa duration and MLAT III correlate to the same extent with native-like L2 attainment and are stronger predictors for phonetic aptitude than other variables.

The results of this study demonstrate that there are different factors which contribute to L2 accent and English aptitude scores. It was observed that AoA also plays some role in L2 phonological acquisition and that "ultimate L2 attainment generally deteriorates with increasing AO" (Jia & Fuse, 2007; Krashen, Long, & Scarcella, 1979). The correlation between length of schwa pronunciation and age of onset of language showed a positive effect of $r = .41$, p (one-tailed) < .05, and the correlation between age of onset of language and English pronunciation scores showed a positive effect of $r = .42$, p (two-tailed) < .05. These findings show that an early exposure to an L2 results in a more native-like L2 accent (here shorter schwa pronunciation) and a more native-like performance, and vice versa. Thus, our data confirms Granena and Long's (2012) claim that "[a]ge of first meaningful second language (L2) exposure, or age of onset (AO), is widely recognized as an indicator of success in second language acquisition (SLA)". However, our results suggest, that it is only one of the indicators amongst others.

Another factor which can explain proficiency in L2 acquisition is the influence of native language phonetic categories on L2 perception. This is due to the fact that more experience with a native language results in the cementing of already built language categories. As the perceptual assimilation model (NLM) by Kuhl, Williams, Lacerda, Stevens, and Lindblom (1992) suggests, the phonetic categories of L1 play an important role in L2 perception and "by adulthood, linguistic experience has had a profound effect on speech perception" (Kuhl et al., 1992, p. 606). Because "[e]xposure to a specific language results in a reduction in the ability to perceive differences between speech sounds that do not differentiate between word [sic] in one's native language" (Goto, 1971; Miyawaki et al., 1975; Strange & Dittmann, 1984; Werker & Tees, 1984; Werker & Lalonde, 1988). Higher native-like pronunciation in early AoA learners assumedly occurs due to the flexibility of speech muscles and articulators in adapting to novel sound productions.

The present data showed an impact of the Persian native vowel system on schwa production. The results of the vowel length measurements revealed a significant difference between the high ability group and the rest of the participants (low ability and medium ability groups). Talented subjects pronounced the schwa significantly shorter than other groups, i.e. around half of the duration of the other ability groups whose schwa duration ranged from 52 to 127 ms. The most significant correlation of our data concerns the correlation between the English pronunciation score and schwa length production in all ability groups $r = -.8$.

For the present study, we used L0 measurements (Hindi imitation task) used by Reiterer et al. (2011 & 2013) as an indication of phonemic coding ability to investigate phonetic imitation and English native-likeness in Persian subjects. We found that ESL learners with better L0 mimicking ability had better English pronunciation performance (meaning shorter schwa production) $r = -.35$, and also there was a significant relationship between English pronunciation and Hindi imitation scores ($r = .36$; p (two- tailed) < .05).

Following Kuhl, Williams, Lacerda, Stevens, and Lindblom (1992) and the principles of the magnet effect, one of the significant findings of our study is that adult L2 learners in their L2 sound production (here vowel production) only partially reproduce some of the vowels which perceptually match their native vowel systems. In our subjects, the longer pronunciation of schwa matched the Persian long vowels /a/ and /e/, but higher ability subjects produced shorter schwas (not represented in the Persian system), hence there is large individual variation in L2 speech production.

As an L0 recognition task, we followed Munson, Edwards, and Beckman (2005) who suggested that novel-word repetition in non-word repetition tasks could be a reliable predictor for L2 learning capacity. Here, our subjects were not supposed to repeat the words but to recognize them in the second phase of the test. We integrated the Llama_D test as this test uses invented non-words and is draws on the subjects' immediate and long-term memory (long-term memory in the sense that the recognition of non-word prompts takes place after a set of stimuli were presented). In this sense it is not a classical WM test. The scores in the Llama_D test revealed a significant, positive relationship with the subjects' English pronunciation score (r = .66) which proves that subjects with better short and long-term memories could achieve a higher native-like attainment in the pronunciation of a second language.

One of the major hypotheses (H6) of our study was concerned with whether WM capacity has a relationship to L2 aptitude. This was corroborated. The results of WM scores with different L2 aptitude tests showed a general, positive correlation. The most significant result was obtained between WM and English pronunciation score (r=. 79). This strong correlation can be interpreted on the basis of Robinson's (2005) L2 aptitude model which focuses on cognitive abilities and suggests that memory plays the most important role in L2 learning success. Our results also support Miyake and Friedman (1998) who considered WM as a fundamental constituent of language aptitude. As mentioned by Rota and Reiterer (2009), a higher working memory capacity is an indispensable criterion for academic achievement.

We observed further strong correlations of WM and language aptitude subtests: WM and MLAT III (r = .63) followed by WM and schwa length pronunciation (phonetic score) with a negative correlation (r = −.61), suggesting that participants with better working memory could perform better in L2 pronunciation. This observation implies that Persian ESL learners with better memory can create ad hoc phonetic category for imitating an English schwa that results in native like, i.e. shorter, pronunciation of the schwa sound.

The next significant result was observed between WM and the MLAT V, which measures retention by means of testing paired associates. This result was almost at the same strength as the correlation between WM and schwa length pronunciation (r = .60) and with MLAT IV (r = .55). This observation indicates that phonemic coding ability, memory and linguistic abilities are similarly fueled by WM.

The result of our Persian near-native L2 speakers (high ability group) supports the work by Abrahamson and Hyltenstam (2008) who concluded that language aptitude can compensate for the later age of onset. In all three English pronunciation ability groups, scores in English aptitude tests correlated highly with English pronunciation scores, and there was a tendency for younger age of onset learners to shorten the schwa duration. Following Wong and Ettlinger (2011) for our language talent measurement, we relied on cognitive aspects of L2 aptitude (other factors being neurophysiological and neuroanatomical aspects). The results of language aptitude tests of this study showed that these measurements are reliable predictors for L2 success or failure of English phonological behavior.

Robinson (2005) argued that these tests are no valid criteria for L2 aptitude measurement in advanced L2 learners. He mentioned that "[b]y 1990, there was also concern that whereas traditional tests such as MLAT were effective in predicting initial progress in language learning, they were seen to be less effective at predicting success at more advanced stages". However, as it was observed in our data, the MLAT battery can be a relatively good predictor for measuring talent in foreign or second language for both early and advanced language learners because of the high correlation of the MLAT battery test results with the English pronunciation scores. The correlation of English pronunciation scores in extreme groups and the MLAT III was $r = .85$. Finally, English pronunciation scores correlated highly with the MLAT V $r = .64$ in all ability groups. These results tentatively prove the validity of the MLAT aptitude tests for predicting L2 learning ability, but we also have to carefully consider that the test battery itself uses English language material and thus contaminates our results by mixing aptitude with proficiency.

Our data suggests that memory capacity is one of the most significant factors in achieving native-like English pronunciation and is highly variable between individuals. For example, the L0 imitation task which is aimed at predicting individuals' abilities in novel sound repetition relies mostly on cognitive abilities because memory ability plays a decisive role in perception and production of L2 sounds. According to Munson, Kurtz, and Windsor (2005), "nonword repetition relies on a number of cognitive processes, such as perceiving and discriminating the acoustic signal, matching the signal with phonological representations in memory, planning the articulatory movements required to replicate the nonword, and executing the response". This assumption maintains that the quality of linguistic input and output is dependent on cognitive abilities.

8 Conclusion

Our experiment shows that the large individual variability in the phonetic L2 success of adults can be attributed to scores in language aptitude, cognitive ability (e.g. memory), AoA, multilingualism/polyglotism and speech imitation ability. Following Rota and Reiterer (2009) in considering higher WM capacity as an indispensable criterion for academic achievement, our data also confirmed that higher language aptitude is reflected by better scores in WM ability tests. It can be seen that individuals with higher WM are better language learners and are able to achieve a more native-like English accent. Better WM capacity is a facilitating factor for L2 learning in that L2 phonological input can be more effectively processed for L2 production and is less reliant on L1 phonological categories (less transfer). This finding leads us to the information processing in individuals and the fact that the difference in ESL learners' linguistic behavior is due to their internal cognitive difference in dealing with linguistic data. The question of whether mispronunciations of ESL speakers are a result of the difficulty of perceiving the exact L2 sound segments or the inability to reproduce novel sounds arises. This argument can be attributed to prototypical perception since, after a certain age, individuals cannot distinguish between words that do not differ in meaning in their mother tongue (Kuhl et al., 1992). Thus, for linguistic production it is important how individuals retrieve the novel linguistic information. The data suggest that ESL learners with better sound imitation possess a better memory and that the quality of the L2 sound segments which they store is higher. Therefore, it is of crucial importance how precisely one perceives and recalls linguistic input. In other words, the higher quality of sound segments in L2 pronunciation is based on the exact phonological information and better processing and encoding of language tokens. Accordingly, the quality of the traces a person stores is reflected in their *emergent* linguistic productions. It can be concluded that the ultimate attainment in L2 production depends on the ability of the brain to categorize, memorize and create/instantiate linguistic input and output at the spur of the moment. This is more attributed to general cognitive and memory ability and the concept of *Emergent Phonology* as a perception/production model (Lindblom, 2000).

Our data additionally shows an influence of AoA and number of languages on ultimate L2 pronunciation proficiency. Younger foreign language learners were more successful in their native-like attainment which suggests that an earlier AoA inhibits more L1 transfer. Thus, in older L2 learners, more experience with a native language is a hindering effect in L2 native-like attainment. It can be concluded that bilingual articulators are more trained to adapt to new speech sound categories, because for imitating a new language sound, they can adapt their speech organs to produce novel sounds. Another observation was that the female subjects of our study performed better than males on all levels. However, the age data showed that females were on average 2.5 years younger than male subjects. Therefore, one way of interpreting the better performance of our female subjects could be their younger age, and hence a better neuromuscular plasticity/flexibility and L2 phonological processing. All in all, more experience with L2 learning, here earlier age of onset of acquisition, and polyglotism are crucial factors in successful foreign language learning as well. Reaching an auditory target is more likely to be achieved by people who are speakers of multiple foreign languages (polyglots). This finding can be compared with the view of Edwards, Beckman, and Munson (2004) that a better phonological processing is a consequence of a larger vocabulary size which can be attributed to multilinguality. In particular, our data posits new evidence for the predictive power of vowel duration (schwa) as a phonetic marker of L2 learning ability.

A full investigation of all language aptitude tests and their influence on L2 phonological achievement was beyond the scope of this paper, so we focused on a few selected aptitude and cognitive ability tests. In short, the most significant correlations of our study involved English pronunciation scores and schwa length pronunciation ($r = -.8$), English pronunciation scores and MLAT III scores ($r = .8$), English pronunciation and working memory scores ($r = .78$), the number of foreign languages spoken and schwa duration ($r = -.74$) and English pronunciation scores and the MLAT V ($r = .64$). Taken together, our results provide evidence of the decisive impact of pronunciation aptitude, phonetic aptitude, phonemic awareness, age of onset, polyglotism and working memory capacity on individual differences in native-like L2 attainment.

References

Abrahamson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition, 30*(4), 481–509. https://doi.org/10.1017/S027226310808073X

Ausubel, D. P. (1963). The psychology of meaningful verbal learning. In *Grune & Stratten*. New York.

Cabrelli Amaro, J., & Rothman, J. (2010). On L3 acquisition and phonological permeability: A new test case for debates on the mental representation of non-native phonological systems. *International Review of Applied Linguistics in Language Teaching, 48*(2–3), 275–296. https://doi.org/10.1515/iral.2010.012

Carroll, J. B. (1981). Twenty-five years of research on foreign language aptitude. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 25–53). Rowley, MA: Newbury House.

Carroll, J. B., & Sapon, S. M. (1959). *The modern languages aptitude test*. San Antonio. Psychological Corporation.

Dörnyei, Z. (2005). *The psychology of the language learner:Individual differences in second language acquisition*. New York: Lawrence Erlbaum.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and Phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language, and Hearing Research, 47*(2), 421–436. https://doi.org/10.1044/1092-4388(2004/034

Erard, M. (2012). *Babel no more: The search for the world's most extraordinary language learners*. New York: Free press.

Goad, H., & White, L. (2006). Ultimate attainment in interlanguage grammars: A prosodie approach. *Second Language Research, 22*(3), 243–268.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia, 9*, 317–323.

Granena, G., & Long, M. H. (2012). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research, 29*(3), 311–343.

Ioup, G. (2008). Exploring the role of age in the acquisition of a second language phonology. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 41–57). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native mandarin speaking children and adolescents: Age-related differences. *Journal of Speech, Language and Hearing Research, 50*(5), 1280–1299. https://doi.org/10.1044/1092-4388(2007/090

Krashen, S. D., Long, M. H., & Scarcella, R. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly, 13*, 573–582.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*(5044), 606–608.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.

Lindblom, B. (2000). Emergent phonology. In Proceedings of the 25th annual meeting of the Berkeley linguistics society, pp. 195–209.
Major, R. C. (2008). Transfer in second language phonology. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 63–83). Amsterdam, The Netherlands: John Benjamins.
Miyake, A., & Friedman, D. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Erlbaum.
Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*, 331–340.
Munson, B., Edwards, J., & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech Language and Hearing Research, 48*(1), 61–78. https://doi.org/10.1044/1092-4388(2005/006
Munson, B., Kurtz, B. A., & Windsor, J. (2005). The influence of vocabulary size, phonotactic probability, and wordlikeness on nonword repetitions of children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 48*(5), 1033–1047. https://doi.org/10.1044/1092-4388(2005/072
Nauchi, A., & Sakai, K. L. (2009). Greater leftward lateralization of the inferior frontal gyrus in second language learners with higher syntactic abilities. *Human Brain Mapping, 30*(11), 3625–3635. https://doi.org/10.1002/hbm.20790
Neuner, S. (2002). Lernen als Verarbeitung von Informationen, Konditionstheoretische Übelegungen zum Fremsprachenunterricht. Babylonia, 10, 35–39.
Obler, L. K., & Fein, D. (1988). *The exceptional brain: Neuropsychology of talent and special abilities*. New York: Guilford Press.
Osgood, C. A. (1946). Meaningful similarity and interference in leaning. *Journal of Experimental Psychology, 36*, 277–301.
Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.
Pimsleur, P. (1966). *The Pimsleur language aptitude battery*. New York: Harcourt Brace Jovanovic.
Rees, J. (2000). Predicting the future of foreign language aptitude. In S. Cornwell & P. Robinson (Eds.), *Individual differences in foreign language learning: Effects of aptitude, intelligence, and motivation* (pp. 187–197). Tokyo, Japan: Aoyama Gakuin University.
Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., et al. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: Functional neuro-imaging and brain morphology. *Frontiers in Psychology, 2*(271), 1–12. https://doi.org/10.3389/fpsyg.2011.00271
Reiterer, S. M., Hu, X., Sumathi, T. S., & Singh, N. C. (2013). Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in Psychology, 4*(782), 1–13. https://doi.org/10.3389/fpsyg.2011.00271
Robinson, P. (2002). *Individual differences and instructed language learning*. Philadelphia, PA: John Benjamins.
Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73. https://doi.org/10.1017/S0267190505000036
Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, 7*, 532–547.
Rota, G., & Reiterer, S. M. (2009). Cognitive aspects of pronunciation talent. In G. Dogil & S. Reiterer (Eds.), *Language talent and brain activity* (pp. 67–96). Berlin: Mouton de Gruyter.
Scovel, T. (1998). *A time to speak: A psycholinguistic enquiry into the critical period for human speech*. Cambridge, MA: Newbury House.
Selinker, L. (1972). *Interlanguage. International review of applied linguistics* (Vol. 10, pp. 209–231).
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed Language Learning* (pp. 69–93). Amsterdam, The Netherlands: John Benjamins.
Skehan, P. (2011). Language aptitude. In S. Gassand & A. Mackay (Eds.), *The Routledge handbook of second language acquisition*. London: Routledge (in press).
Spolsky, B. (1995). Prognostication and language aptitude testing, 1925–1962. *Language Testing, 12*(3), 321–340.
Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics, 36*, 131–145.
Tewes, U. (1994). *Hamburg-Wechsler-Intelligenz-Test für Erwachsene Revision*. Bern, Switzerland: HAWIE-R.
Thomason, S. G. (2001). *Language contact*. Edinburgh, UK: Edinburgh University Press.
Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly, 4*(2), 123–130.
Wen, Z. (2011). Foreign language aptitude. *ELT Journal, 66*(2), 233–235. https://doi.org/10.1093/elt/ccr068
Werker, J. F., & Lalonde, C. E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychobiology, 24*, 672–683.
Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49–63.
Wong, P. C., & Ettlinger, M. (2011). Predictors of spoken language learning. *Journal of Communication Disorders, 44*(5), 564–567.