Second Language Learning and Teaching

Ewa Waniek-Klimczak
Linda R. Shockey *Editors*

Teaching and Researching English Accents in Native and Non-native Speakers

Springer

Second Language Learning and Teaching

Series Editor

Mirosław Pawlak

For further volumes:
http://www.springer.com/series/10129

About the Series

The series brings together volumes dealing with different aspects of learning and teaching second and foreign languages. The titles included are both monographs and edited collections focusing on a variety of topics ranging from the processes underlying second language acquisition, through various aspects of language learning in instructed and non-instructed settings, to different facets of the teaching process, including syllabus choice, materials design, classroom practices and evaluation. The publications reflect state-of-the-art developments in those areas, they adopt a wide range of theoretical perspectives and follow diverse research paradigms. The intended audience are all those who are interested in naturalistic and classroom second language acquisition, including researchers, methodologists, curriculum and materials designers, teachers and undergraduate and graduate students undertaking empirical investigations of how second languages are learnt and taught.

Ewa Waniek-Klimczak
Linda R. Shockey
Editors

Teaching and Researching English Accents in Native and Non-native Speakers

Springer

Editors
Ewa Waniek-Klimczak
Department of English Language
and Applied Linguistics
University of Lodz
Lodz
Poland

Linda R. Shockey
Department of English Language
and Literature
University of Reading
Reading
UK

ISSN 2193-7648 ISSN 2193-7656 (electronic)
ISBN 978-3-642-24018-8 ISBN 978-3-642-24019-5 (eBook)
DOI 10.1007/978-3-642-24019-5
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012952734

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Extending the scope of interest beyond the description of individual varieties of English, the broad 'accents' theme as developed in this volume, includes phonetic and phonological characteristics of first, second or third languages. Contributors to the volume teach and do research on a variety of relevant aspects, as represented in our section headings.

The multi-perspective approaches offered in this volume all develop one theme: variability in pronunciation. We believe the book offers a unique combination of psycholinguistic, sociolinguistic and pedagogical approaches, with individual contributions investigating the effect of selected conditioning factors on the pronunciation of English and other languages. Despite the wealth of approaches, a strong phonetic background unifies individual contributions to the volume. The book contains a large body of original, primary research which will be of interest to experienced scientists, practitioners and lecturers as well as to graduate students planning to embark on empirical investigation of the nature of a sound system.

Ewa Waniek-Klimczak
Linda R. Shockey

Contents

Part I
Pronunciation: Production and Perception

Teachability and Learnability of English Pronunciation Features for Vietnamese-Speaking Learners

Una Cunningham

Abstract Anyone who has tried to learn a language with a very different sound system will understand the challenges faced by speakers of a language as different as Vietnamese who are attempting to learn to speak English in a way that is intelligible to non-speakers of Vietnamese. Many learners have very limited opportunity to hear model pronunciations other than their teacher's, and no opportunity at all to speak in English outside the classroom. Vietnamese-accented English is characterised by a number of features which ride roughshod over English morphosyntax, resulting in speech that is extremely difficult to reconstruct for the non-Vietnamese-speaking listener. Some of these features appear to be more difficult to learn to avoid than others. Phonotactic constraints in L1 appear to be persistent even in L2, and L1 phonological rules will, apparently, often apply in L2 unless they are blocked in some way. Perception of salient (to native listeners) target pronunciations is often lacking, and learners may not be aware that their pronunciation is not intelligible. Despite years of language study, many learners are unable to produce some native speaker targets. Vietnamese learners typically exhibit a set of characteristic pronunciation features in English, and the aim of this study is to see which of these are susceptible to remediation through explicit teaching. This explicit teaching is compared with a less direct, less interactive kind of teaching, involving drawing native and native-like pronunciation of problematic features of English pronunciation to the learners' attention. The results of this study can then be interpreted in terms of teachability and learnability, which do not always go hand in hand. If we understand what kinds of phonetic features are teachable and how learnability varies for different features, we can target those features where there is a good return for effort spent, resulting in efficient teaching.

U. Cunningham (✉)
University of Stockholm, Stockholm, Sweden
e-mail: una.cunningham@isd.su.se

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_1, © Springer-Verlag Berlin Heidelberg 2013

1 Introduction

One of the most mind-expanding experiences I have had in recent years has been attempting to learn Vietnamese at an age well past any critical period. I found that I was able to observe my pronunciation learning attempts—strategies I used to try to remember the exact quality of the high and mid-high back unrounded vowels and diphthong, and particularly how I tried to come to terms with the six tones used in Hanoi. The first challenge, as a first-time learner of a tone language is to notice the tones. Then comes a period of trying to distinguish them reliably from each other. The final challenge is to be able to produce tone in a way that is interpretable to listeners, first in isolated words, then in running speech.

The typological distance from English to Vietnamese is approximately equal to that between Vietnamese and English, and it is reasonable to assume that the kinds of difficulties I experienced in my attempts to pronounce Vietnamese are matched by different but comparable difficulties which face Vietnamese-speaking learners of English. Students have been heard to complain that it is difficult to remember how to pronounce words without tones, and the English vowel system offers a real challenge. English phonotactics is daunting and the semantic weight attached to the tongue-twisting clusters that burden every other English syllable sets the odds against success.

Vietnamese learners of English typically have a much tougher time with pronunciation than with syntactic or lexical acquisition. Vietnamese-accented speech is often difficult for non-speakers of Vietnamese to understand, even if the speaker can communicate well in written English. There are a number of reasons why this is so, and these have been the topic of previous work (Cunningham 2009b, c). The interesting observation that a non-native accent is generally more easily understood by other speakers of the same accent has been documented, e.g. by Jenkins (2002).

The fact that English pronunciation is not tested in the important national university entrance exams in Vietnam has, of course, a knock-on effect such that little effort is put into pronunciation in Vietnamese schools. On the other hand, the difficulty of English pronunciation for Vietnamese learners is legendary, and there is a feeling that the rewards of effort spent in the pursuit of better pronunciation are miserly. These conditions raise the twin questions of just how teachable and just how learnable English pronunciation is for these learners.

2 Teachability of Pronunciation

The concept of teachability of language was developed by Pienemann who wrote about it in connection with his Teachability hypothesis, suggesting that specific syntactical constructions could only be learned after mastery of parts of language that come earlier in second language development (Pienemann 1989). He also

found that instruction has an accelerating effect on acquisition for learners who are ready for it (Pienemann 1984). This is not to be confused with the kind of teachability referred to by Tubbs who saw it as a property of the student rather than of the material to be learned (Tubbs 2003).

If, as has been suggested e.g. by Taylor (1993) and Jenkins (2000), Pienemann's model can be applied to pronunciation, it would suggest that while some undesirable features of L1 accents can be remedied informally, others are not affected by explicit teaching

Further, this approach implies that there is an order of acquisition of pronunciation features for second language learners, just as has been found for child language acquisition (Jakobson's 1942 order of acquisition steps). Examples of this are Demuth and McCullough (2009) and Kirk and Demuth (2005) working on the acquisition of clusters in French and English speaking children and Shillcock and Westermann (1997) on phonotactics, though Menn and Stoel-Gammon (1995) found inconsistencies in the order of acquisition of phonemes and allophones, and Dutch children have been found to vary in whether they acquire word-initial or word-final clusters first.

Factors other than the order of acquisition of pronunciation features may also affect the teachability of a particular pronunciation feature. The availability of pronunciation models is one such factor (Cunningham 2009a). Even with widespread access to the Internet, many Vietnamese learners do not regularly listen to English spoken by others than their teacher and classmates. Teachability can only be considered within a given educational context. The topic of this paper is the teachability of particular features of English pronunciation (as specified in targets relative to a particular pronunciation model) in a given educational setting with a given configuration of teachers and learners. It might be possible to alter the outcome of teaching by changing one or more of these parameters, such as the model, or the students' access to the model. The way the teaching is carried out may also be relevant.

After a recent conference in Ottowa, Derwing and Munro note that little work has, until recently, been oriented towards an examination of the efficacy of pronunciation teaching (Derwing Tracey and Munro Murray 2010). Traditionally research on L2 phonological acquisition has been concerned with learners errors, but Derwing and Munro point to more recent work which has been looking at the relationship between accent and intelligibility, cognitive processes underlying phonological learning, the evaluation of L2 speech, prosodic influences on the perception of accented speech, and the role of ethnic affiliation and identity in L2 speakers' oral production.

Celce-Murcia et al. (2010) list a number of ways pronunciation has been taught, with an emphasis on word-level accuracy. Pronunciation teaching can potentially involve many kinds of activities, such as:

- lectures and exercises on the phonetics of the model variety,
- extensive listening activities (listening for content),
- extensive speaking activities in a communicative learning context,

- real communication outside the classroom (with or without feedback),
- intensive listening activities (listening for form) and noticing activities,
- listen and repeat activities with different kind of feedback:
  - student listening in real time or to recording of own attempt,
  - comment from peer or teacher, correction from peer of teacher,
  - remodelling from teacher or recorded material,
  - visual feedback (CALL),
  - speech recognition feedback (peer, teacher or CALL).

It is extremely difficult to compare and evaluate the efficacy of these kinds of teaching. Typically learners will be exposed to several of these at a time, as well as other kinds of language work that are not specifically aimed at improving pronunciation, but which may well have an effect. The importance of interaction and feedback in the learning of pronunciation is often assumed, but there is little evidence to support this, and it is very difficult to measure the effect of this kind of input. Literature in this area is skimpy, and in an attempt to explore this further I have initiated a series of interviews with L2 speakers of English who have either worked with their accents or who have achieved near-nativeness in some respect. Results will be reported in future publications.

Might the persistence of some of the non-English-like features of the English produced by Vietnamese learners who are competent users of written English suggest that these pronunciation features are being taught in the wrong order? Or might the fact that even very advanced users of English fail to master some features of English pronunciation actually point to them not being learnable for these learners? Of course there are Vietnamese learners of English who do learn to pronounce English without the characteristic features associated with Vietnamese accents, but they may be applying other kinds of learning strategies than the usual ones.

3 Learnability of Pronunciation

There is no doubt that there are individuals who are more successful than others at achieving their targets for the learning of pronunciation. Some individuals acquire native-like pronunciation despite a late start and limited opportunities for interaction with native speakers of the target language. In at least some such cases the result is less due to unusual talent or aptitude than to hard work. Motivation and identity are also important factors in the equation. Some individuals feel that attempting to acquire a native-like accent is a rejection of their own culture which endangers their personal identity.

Some learners, perhaps particularly where there is a local variety of the target language, such as is the case for speakers of English in Nigeria or Pakistan, feel strongly that the local variety is the only legitimate target for their pronunciation, regardless of any problems this might cause when using English with speakers

from outside this community. Others, with less personal identification or simply with less phonetic sensitivity, feel that a native-like accent is the obvious target for an ambitious language learner.

Jenkins, in her Lingua Franca core (Jenkins 2002), identified some features of English pronunciation as being more worthy than others of attention, at least if the target is to be intelligible to as wide a variety of native and non-native speakers of English as possible. For example, one of the features she mentioned as not being important for intelligibility and yet very difficult to acquire is the English inter-dental fricatives. Others have since then looked at various ways of ranking pronunciation targets. Scales et al., for example looked at learners' preferences, and found that learners preferred more intelligible accents (Scales et al. 2006). Others (unsurprisingly) have also found that listeners prefer speech that is intelligible to them (Burda and Hageman 2005; Kennedy and Trofimovich 2008; Munro and Derwing 1995; Volberg et al. 2006). Cunningham-Andersson and Engstrand isolated and elicited reaction to various features of Finnish-accented Swedish (Cunningham-Andersson and Engstrand 1989) and could rank individual features and combinations of features for the disturbance associated with them. But these rankings do not concern themselves with how possible it is for a learner to achieve the target.

It is generally held that repeated exposure is a significant factor in language learning (Cervantes and Gainer 1992; Flege et al. 1997). Learners need to meet new forms, be they grammatical structures, items of vocabulary or new sounds, in a number of contexts (Saragi et al. 1978) before they can be said to have learned them. The approach used in this study is an attempt to optimise the exposure of students to a limited set of pronunciation challenges faced specifically by Vietnamese learners of English. However important acquisition of a feature might be, some features seem to be particularly difficult for speakers of a given language to acquire.

4 Vietnamese-Accented English

Recent work by Cunningham outlines some of the more problematic issues for intelligibility Cunningham (2009a, b, c) in the speech of Vietnamese students of English. There are many segmental and suprasemental differences between the two languages, with English suprasegmentals being especially challenging for Vietnamese-speaking learners.

4.1 Method

This study is set in Hanoi College of Business and Tourism, a vocational further education college which has some 400 students per year. Their days are filled with timetabled teaching from morning to late afternoon, and they have very little time

for unscheduled activities, even their own reading and listening work. The college language laboratory is not in use due to lack of funds for maintenance. Otherwise the classrooms in the college are not well suited to listening work. The sound environment in classrooms is quite poor as the doors and windows are generally open, to provide better airflow. There are fans rotating and the noise of the Hanoi traffic pervades everything. The classrooms have hard surfaces which reflect sound, and there are around 60 students in a class. Teachers generally use microphones to make themselves heard, and there is competition from the students' own conversation and mobile phone activity.

Though throughout Vietnam little emphasis is put on oral skills, at this particular Further education college, teaching staff and the college leadership are well aware of the importance of oral proficiency, perhaps particularly in the fields of business and tourism.

The author was invited to develop and deliver a short course in English pronunciation for 110 students of tourism in four classes at a College of Business and Tourism in Hanoi. Two courses were developed: first, a short course in pronunciation by listening, packaged as nine audio files was developed for three small classes (49 students in all). Secondly, an eight-lesson classroom course in pronunciation was delivered to 61 students in a single class. Specific pronunciation features are in focus in each session, dealing with specific needs of Vietnamese learners as established in earlier work (Cunningham 2009a, b). The focus of both the classroom and the audio file teaching is intelligibility. Students in Vietnam are not generally served by an aspiration to approach native-like pronunciation. The set of prioritised pronunciation targets for English as an International language proposed by Jenkins (2002, 2005) are broadly compatible with the features selected for the courses in this study.

The taught course was organised in eight sessions (two or three a week for three weeks). Each session concentrated on a particular aspect of the pronunciation of English, such as vowel quality, vowel duration, consonant articulation, final consonants, final clusters and initial clusters. A typical session began with general chitchat about the weather or the traffic or what had been going on since the last class, to allow the students to hear and get used to the speech of the teacher and for the braver students to interact with the teacher. Then the topic for the day's class was introduced and modelled for the students to hear. Contrasts between sounds were demonstrated in isolation and in running speech, and the students were asked to try to discriminate between contrasting sounds and words (e.g. *nine, line, Nile),* firstly in the teacher's speech, and then in their partner's speech. Students worked in pairs, attempting to produce the target sounds initially in a way such that their partner could reliably know which of two or three contrasting sentences they were reading. Finally, groups of students were each asked to nominate a speaker to produce the contrasting sentences for the teacher to distinguish.

The remaining students were each given an mp3-player with nine short lessons recorded on it. The lessons covered the same points as the taught sessions. Each lesson had a text for the students to listen to, comments from the teacher, and then

excerpts from the text illustrating the use of the target sounds which were the topic for that session.

Before and after the intervention, each student in both groups was recorded by an assistant reading a text, sentences and words designed to elicit the features which are focussed on in the classroom and podcasted courses, such as vowel quality and duration, initial and final consonants. The sentences were:

1. *Please sit on this little white seat. These wide black seats are still wet.*
2. *Can you ask his sister not to hiss and scream if she scolds him when he's late?*
3. *Bring me five sheep skins and four goat skins after lunch.*
4. *She can see him with sixteen sheep not six tin ships.*

In addition, the groups were given an opportunity to give written feedback about the courses, and this was complemented by interviews with teachers on their view of the course and the material. The ten phonetic features in focus for the teaching material were selected on the basis of previous work on the intelligibility of the English spoken by Vietnamese speakers (Cunningham 2009c). They include some of the most problematic characteristic features of the English spoken by Vietnamese-speaking learners:

1. distinction between the vowels of *ship* and *sheep,*
2. the pronunciation of dental fricatives,
3. final clusters (regardless of morpheme boundaries, e.g. *seats, ask, six*),
4. initial glides (*wet, wide, one*),
5. distinction between/n/and/l/(*lunch, school*),
6. temporal, glottal or fortis/lenis distinction to mark postvocalic consonant voicing (*his-hiss, white-wide*),
7. initial clusters (*scream, scold*),
8. elision of the coda (*wet, late*),
9. distinction between the first sounds in *see* and *she,*
10. distinction between/f/and/p/(e.g. in *if* and *after*).

For each speaker, a point was given for each of the features pronounced in a target-like way, so that a speaker could score a maximum of ten points for each recording. The results for the group before and after the intervention were tested using two-tailed *t* tests.

4.2 Results

In the taught course the mean score for the 10 pronunciation features for the students before the teaching began was 4.24 (n = 50) which means that they on average successfully pronounced 4.24 of the 10 target pronunciations. In the mp3-group the average score before the students received the mp3-players was marginally higher 4.77 (n = 43). A two-tailed student's *t* test shows that this

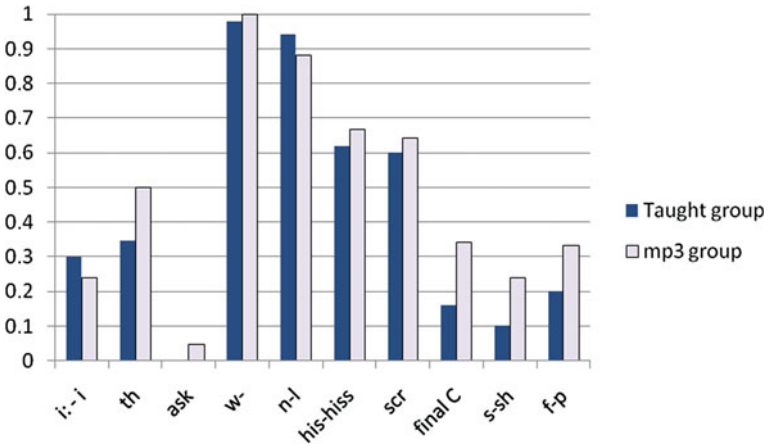


Fig. 1 Comparison between the groups at the beginning of the study broken down by the ten phonetic features which were observed in the students' pronunciation and focused on in the teaching sessions and podcasts

difference between the classes before the teaching began was weakly significant ($p(t) = 0.047$).

The taught class includes 61 students. Fifty of these students chose to participate in the study after full information was given to them, and 51 took part in the final recording. In the mp3-group, 42 took part in the initial recording and only 32 of these took part in the final recording. Figure 1 shows the breakdown for the comparison between the two groups before the start of the intervention by pronunciation feature.

Notice that there is not much difference between the groups. The only feature difference that is significant (at the 5 % level) is feature 8, the pronunciation of the final stop in words like *wet*, where the mp3 group had, on average, a more target-like pronunciation, with a clear stop being made more often.

Figure 2 shows a breakdown of the results before and after the eight teaching sessions. Notice the large improvement in the ability of students in six of the ten observed points (features marked * show improvements significant at the 5 % level and those marked ** are significant at the 1 % level): to distinguish between the vowels in *seat, sheep* and *sit, ship* (feature 1**), to pronounce interdental fricatives (feature 2*), to make postvocalic voicing distinctions (feature 6*), to pronounce a clear/t/at the end of words like *wet* and *seat* (feature 8*), to distinguish between the initial sounds in *see* and *she* (feature 9**) and to pronounce final/f/as [f] rather than [p] (feature 10**). Of the other four features studied, the students had not much difficulty with features 4 and 5 already at the start of the study, while they still could not master final clusters (feature 3) at all by the end of the study, and there was no improvement in their ability to pronounce initial clusters at the end of the study.

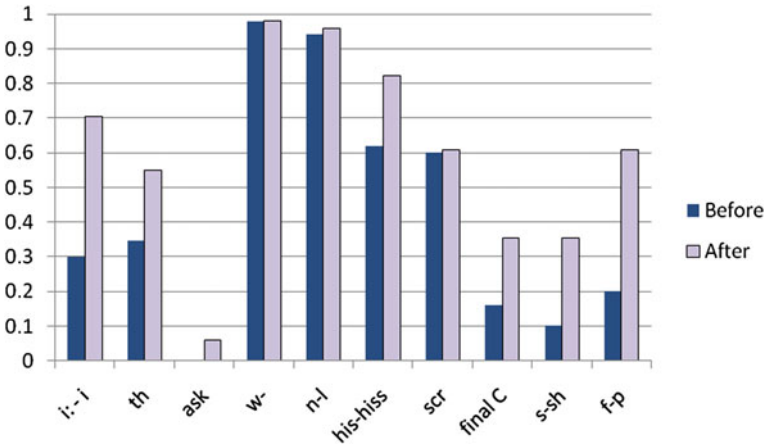


Fig. 2 Comparison for the taught group before and after the study, broken down by the ten phonetic features which were observed in the students' pronunciation and focused on in the teaching sessions

In no case is there a significant difference for the group which was given the mp3 players with the prerecorded audio lessons. It has to be concluded that the mp3 group made no progress whatsoever as measured by this study.

5 Discussion

These students have had very little opportunity to speak English with non-Vietnamese speakers and for most of them this was their first close encounter with a "westerner". Thus the teaching sessions meant a huge increase to their exposure to a native speaker of English with all that entails in terms of quantity of speech material they hear and are required to engage with, and the rigorous requirements placed on their speech when they were required to speak in class in such a way that a native-speaking interlocutor could understand their speech. This required many repetitions and in many cases the students had to resort to writing down what they wanted to say. This was, of course, a disappointment to them and their teacher. However, the taught group appears to have benefitted significantly from the eight teaching sessions they received.

In terms of the teachability and learnability of individual pronunciation features, it seems that consonant clusters, both initial and final, are particularly resistant to learning. Vietnamese phonotactic constraints mean that there are no clusters in Vietnamese words or loan words, and there are severe limitations on what is admissible in the coda of syllables. The initial cluster production was about 60 % accurate, and there was no change in this during the period of the study. There were very few instances of target-like production of final clusters in either

group of these students either before or after the intervention period. So in Peinemann's terms (op cit), perhaps the students are not ready to learn the production of final clusters at this stage, and thus cannot be helped by teaching. Jenkins (2000, 119) claims that "where the difficulty with an L2 English pronunciation feature is universal", which is perhaps true of the fiendishly complex final clusters offered by languages like English, "we are looking at an item that may well be unteachable". Initial clusters seem to be easier for these learners to acquire than final clusters.

It is interesting to set this finding alongside studies of children's L1 acquisition of clusters. Studies of the L1 acquisition of English, where there is a morphological advantage for early acquisition of final clusters show that final clusters are acquired before initial clusters (Kirk and Demuth 2005). Initial clusters were, however, observed to be acquired before final clusters by French-speaking infants by Demuth and McCullough (Demuth and McCullough 2009). They attribute the fact that complexity appears at the beginning of words before at the ends of words to French not having heavy word-final morphophonology, unlike English. They also point to possible alternative syllable analyses that may leave final consonants in French final clusters in another syllable than a preceding consonant. They further suggest that final clusters are typologically more marked than initial clusters and thus potentially likely to be acquired later.

The features that were most susceptible to change due to the intervention, the difference between *sheep* and *ship*, the difference between /p/ and /f/ and between /ʃ/ and /s/ are perhaps easily demonstrated by a teacher, with appropriate exaggeration of the differences to help students to become aware that there is a difference. The three features where there was a less but still significant difference due to the intervention, interdental fricatives, release of final /t/ and the use of vowel lengthening to signal post vocalic voicing were also fairly easy to demonstrate, and to work with. The remaining two features, the difference between /l/ and /n/ and initial /w/ were not very problematic for most students in this group even before the intervention.

Unfortunately, the mp3 group in this study made no progress. This could have a number of explanations. Firstly, it may be the case that the basic premise of this study is flawed. Practice listening to English (extensive listening) and having features of English pronunciation in use be brought to one's attention (intensive listening) may not actually be helpful to students. Secondly, it may be that the students did not actually listen to the material in the way intended. Thirdly, it may be that the students listened once through the material and then did not listen again, thus not gaining enough exposure to the material to make a difference. This could be because they found the material too difficult or simply uninteresting, or because they were not motivated to spend time on this work. This means that although the study did show a gratifying improvement in the specific pronunciation features studied for the taught group, this study cannot show an effect of listening on pronunciation. Obviously further studies are required here, perhaps with a student group that can be more closely followed and supported for a longer period.

References

Burda, Angela Noel, and Hageman, Carlin F. (2005). Perception of accented speech by residents in assisted-living facilities. *Journal of Medical Speech-Language Pathology, 13*(1), 7–14.

Celce-Murcia, Marianne, Brinton, Donna and Goodwin, Janet M. (2010). *Teaching Pronunciation : A Course Book and Reference Guide* (2 ed.). Cambridge: Cambridge University Press.

Cervantes, Raoul, & Gainer, Glenn. (1992). The effects of syntactic simplification and repetition on listening comprehension. *TESOL Quarterly*, 26(4), 767–770.

Cunningham, Una. (2009a). Models and targets for English pronunciation in Vietnam and Sweden. *Research in Language* 7, 113–128.

Cunningham, Una. (2009b). *Phonetic correlates of unintelligibility in Vietnamese-accented English*. Paper presented at the FONETIK 2009 The XXIIth Swedish Phonetics Conference. Available at http://www2.ling.su.se/fon/fonetik_2009/108%20cunningham_fonetik2009.pdf.

Cunningham, Una. (2009c). Quality, quantity and intelligibility of vowels in Vietnamese-accented English. In Ewa Waniek-Klimczak (Ed.), *Issues in Accents of English II: Variability and Norm*. Newcastle: Cambridge Scholars Publishing Ltd.

Cunningham-Andersson, Una & Engstrand, Olle. (1989). Perceived strength and identity of foreign accent in Swedish. *Phonetica*, 46(4), 138–154.

Demuth, Katherine & McCullough, Elizabeth. (2009). The longitudinal development of clusters in French. *Journal of Child Language*, 36(2), 425–448.

Derwing, Tracey. M., & Munro, Murray. J. (2010). Symposium - Accenting the positive: Directions in pronunciation research. *Language Teaching, 43*(3), 366–368.

Flege, James Emil, Bohn, Ocke-Schwen, & Jang, Sunyoung. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4), 437–470.

Jakobson, Roman. (1942). Kindersprache, Aphasie und allgemeine Lautgesetze. Uppsala: *Uppsala Universitets årsskrift* 9.

Jenkins, Jennifer. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

Jenkins, Jennifer. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, 23(1), 83–103.

Jenkins, Jennifer. (2005). Implementing an international approach to English pronunciation: The role of teacher attitudes and identity. *TESOL Quarterly*, 39(3), 535–543.

Kennedy, Sara & Trofimovich, Pavel. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review-Revue Canadienne Des Langues Vivantes*, 64(3), 459–489.

Kirk, Cecilia, & Demuth, Katherine. (2005). Asymmetries in the acquisition of word-initial and word-final consonant clusters. *Journal of Child Language*, 32(4), 709–734.

Menn, Lise, & Stoel-Gammon, Carol. (1995). Phonological development. In Paul Fletcher & Brian MacWhinney (Eds.), *The Handbook of Child Language*. Oxford: Blackwell.

Munro, Murray. J., & Derwing, Tracey. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of 2nd-language learners. *Language Learning*, 45(1), 73–97.

Pienemann, Manfred. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition*, 6(2), 186–214.

Pienemann, Manfred. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics* 10(1), 52–79.

Saragi, Thomas, Nation, Ian Stephen Paul, & Meister, Gerold Fritz. (1978). Vocabulary learning and reading. *System*, 6(2), 72–78.

Scales, Julie, Wennerstrom, Ann., Richard, Dara, & Wu, Su Hui. (2006). Language learners' perceptions of accent. *TESOL Quarterly*, 40(4), 715–738.

Shillcock, Richard, and Westermann, Gert. (1997) The role of phonotactic range in the order of acquisition of English consonants. In Wolfram Ziegler & Karin Deger (Eds.) *Clinical Linguistics and Phonetics*. London: Whurr.

Taylor, David S. (1993). Intonation and accent in English: What teachers need to know. *International Review of Applied Linguistics in Language Teaching*, 31(1), 425–435.
Tubbs, Nigel (2003). The concept of "teachability". *Educational Theory*, 53(1), 75–91.
Volberg, Leonie, Kulka, Marko, Sust, Charlotte. A., & Lazarus, Hans. (2006). Speech intelligibility and the subjective assessment of speech quality in near real communication conditions. *Acta Acustica United with Acustica*, 92(3), 406–416.

On the Irrelevance of Sounds and Prosody in Foreign-Accented English

Jolanta Szpyra-Kozłowska

Abstract The author argues that current research on phonetic priorities in ELT, with its focus on segments and prosody, is misguided and that emphasis should be shifted towards learners' training in the production of words whose idiosyncratic erroneous rendition does not result from their inability to articulate foreign sounds correctly, but which is caused by various interference factors (e.g. *Disney* pronounced by many Polish learners as [ˈdʲisnej]). It is argued that the use of such severely distorted items (local errors) has grave consequences for linguistic communication, more serious than segmental and suprasegmental inaccuracies (global errors) and should, therefore, be pedagogically prioritized. In order to verify this claim, two experiments have been carried out in which 40 native-speakers of English were asked to assess two phonetic versions of the same passage: one produced by a Polish learner of English with poor segmental and suprasegmental pronunciation, but no major phonological distortions of words, and another recording made by a speaker with the correct rendition of segments and prosodies, but with several seriously mispronounced words, common in Polish English. The assessment concerned the samples' degree of comprehensibility, foreign-accentedness and annoyance for the listeners. The experimental data show that on all three counts the participants' judgements were more severe in the case of the version with local errors than with global errors. The same results were obtained in the second experiment in which the samples' intelligibility was examined in a dictation test.

J. Szpyra-Kozłowska (✉)
Maria Curie-Skłodowska University, Lublin, Poland
e-mail: j.szpyrakozlowska@chello.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_2,

1 Introductory Remarks

In view of various pressures imposed by course requirements and having a limited time for teaching pronunciation, an English teacher faces numerous dilemmas. Which elements of English phonetics should be selected for instruction? Should the focus be on segments or rather on prosody? Which properties of English consonants and vowels, and which prosodic features should be given close attention and which can be neglected with little loss for communication? In brief, the following question inevitably arises: what are the pedagogical priorities in English pronunciation instruction?

Clearly, there is no simple answer to these queries since pedagogical priorities, as argued by Celce-Murcia et al. (2010), can only be established when different variables of the teaching process are taken into consideration, such as its participants and their linguistic background, their specific needs and problems. Nevertheless, as emphasized by many pronunciation specialists (e.g. Kenworthy 1987; Derwing and Munro 2005), they should be set in such a way as to ensure achieving what is the primary goal for the majority of learners, i.e. intelligibility.

A distinction should be made between *basic intelligibility*, which allows for rudimentary communication, but puts a considerable strain on the listener and requires much effort on their part to understand the message and *comfortable intelligibility*, to use Abercrombie's (1949) term, which puts little or no strain on the listener.

While achieving comfortable intelligibility is frequently accepted as the main target of phonetic instruction, there is no agreement as to the role which specific aspects of English pronunciation play in it. According to some researchers, these are mostly segmental errors that impede intelligibility (e.g. Jenkins 2000), while according to others (e.g. Derwing and Munro 1997) these are prosodic inaccuracies that are detrimental to it. In view of such controversies, as argued by Field (2005, 399), "the most pressing issue in L2 pronunciation research today is the quest to identify the factors that most contribute to speaker intelligibility."

In this paper we intend to advance a claim that recent pronunciation research on intelligibility which focuses on the role of various segmental contrasts and different aspects of prosody, i.e. on the so-called *global errors* made by the learners, fails to take into account a particularly important aspect of foreign-accented English, namely that of numerous seriously deformed words the use of which, as we shall argue, frequently hinders successful communication far more than other phonetic errors. Apparently learners of English find many words of this language difficult to remember in their phonetic shape not because they contain problematic sounds or sound sequences, but due to a variety of interference factors (discussed in Sect. 2). To put it differently, the items in question, often referred to as *local errors* (Sobkowiak 1996) or *item errors*, are stored in the learners' phonetic memory with phonologically deviant representations.

For instance, many Polish learners of English mispronounce a phonetically seemingly simple word *foreign* as [fo'rejn] although it contains no particularly difficult segments and has penultimate stress which is typical of Polish. Clearly,

Table 1 Characterizing global and local errors

Global errors	Local errors
Recurring mispronunciations of foreign sounds and prosodies which create a foreign accent and result mainly from L1 phonological and phonetic transfer, e.g. E jazz > PE [džes] E foreign > PE ['for^{j}in]	Idiosyncratic mispronunciations of individual words in which, apart from global errors, there are other phonological and phonetic deviations from the original, due to various interference factors. They are stored in the learner's phonetic memory with the incorrect segmental and/or prosodic structure, e.g. E foreign > PE [fo'rejn], E Disney > PE ['d^{j}isnej]

the digraph <ei> that occurs in spelling suggests to Poles a diphthongal pronunciation (found, for instance, in items with the same sequence of letters, such as *reign* and *feign*) or, to be more exact, a sequence of a vowel and the palatal glide. The ultimate stress placement, alien to Polish, is somewhat more mysterious. It can be attributed to the presence of what is assumed to be a heavy stress-attracting final syllable (as in *insane, polite*) or to analogy with the stress pattern of words with the prefix *for-*, such as *forget, forgive* and *forbid.* The same item, when pronounced without such distortions is rendered in Polish English as ['for^{j}in]. In this version global errors are present, i.e. each English segment is replaced with a corresponding Polish consonant and vowel, with an additional palatalization of the rhotic by the following high front vowel.

Thus, phonetically deviant words which are examined in this paper are characterized by segmental errors in which English phonemes are substituted not by their closest counterparts in the learners' native language (e.g. the replacement of the English glottal fricative with the Polish velar fricative), but by segments which are phonologically and phonetically often remote from them (e.g. English / ɪ / rendered as Polish [ej]). Such substitutions frequently involve more segments than one (as in *Disney* pronounced as [d^{j}isnej]) and may be accompanied with the incorrect placement of stress (e.g. *foreign* realized as [fo'rejn]).

Table 1 sums up the above discussion on the distinction between global and local errors.[1]

While the exact number of local errors is evidently impossible to establish as it largely depends on the learners' level of language proficiency and their individual phonetic aptitude, it is certainly large enough to deserve a prominent place in pronunciation instruction. Sobkowiak (1996) presents about 700 of what he calls "words commonly mispronounced" in Polish English, but this list can easily be extended, particularly if we include numerous proper names like *Disney, Turner* or *Presley*, which are also problematic for foreign learners.

[1] The distinction between global and local errors is not always sharp. It is not clear, for instance, how to classify the overgeneralization that the <ate> sequence is often interpreted as [eit] in nouns, such as *certificate, climate* or *palate.* Such errors are local in the sense that they concern a particular sequence of letters, but are not restricted to a single item.

The importance of local errors lies not only in their large number, but in their high frequency in learners' English, a considerable degree of fossilization and resistance to improvement.

It should be noted that the failure to make a distinction between global and local errors has often some serious consequences and leads to false conclusions. Jenkins (2000), for instance, claims that the quality of vowels is generally irrelevant in her Lingua Franca Core, which is supposed to ensure mutual intelligibility of international users of English, except for the vowel known as long schwa, as in *girl, burn, firm.* As pointed out by Szpyra-Kozłowska (2005), it is fairly surprising why of all English vowels the least frequently occurring monophthong (Cruttenden 1994, 136) should be singled out as crucial for intelligibility. Jenkins provides only two examples in which the mispronunciation of this segment led to communication breakdowns in her data, i.e. *bird* pronounced as *bard* and *birthplace* realized as *bathplace*. It seems that the problem probably does not concern her informants' inability to articulate this particular vowel—no such evidence has been provided, but involves instances of local errors that result from the incorrect encoding of these particular items in the learners' memory. A similar type of mispronunciation can frequently be observed in Polish English, where *Turner* is often pronounced as *Tarner, Murphy* as *Marphy* and *purple* as *parple*, which has nothing to do with the difficulty of producing long schwa by Poles, but follows from an incorrect overgeneralization concerning the realization of the letter <u> as some kind of [a]-like vowel. If this reasoning is correct, international learners of English should be given phonetic training not so much in the production of long schwa as in the pronunciation of problematic words, such as *bird, birth, Turner* or *purple*, possibly coupled with the teacher's comments on some letter-to-sound correspondences. Similarly, Avery and Ehrlich (1992), in their advice to pronunciation instructors, maintain that Polish learners of English should be trained in the production of [v], as they often say *willage* or *walley* instead of *village* and *valley.* This is another obvious instance of confusing local errors with global ones since Poles do not have to practice the labiodental fricative as it is found in hundreds of Polish words. What they do need to practice, however, is to pronounce the words *village* and *valley* with the initial fricatives and not labiovelar glides.

In what follows we are going to argue that local errors, extremely frequent in learners' English and generally disregarded both in teaching materials and in current pronunciation research, pose a considerable threat to various aspects of effective communication and intelligibility in particular. Consequently, they should be regarded as a top pedagogical priority.

2 Sources of Local Errors

The question that arises in connection with the conclusion arrived at in the preceding section is why the serious issue of phonetically distorted words fails to be addressed in phonetic manuals. In order to answer it, it is necessary to examine the major sources of such local errors. Their classification presented below is largely based on Sobkowiak (1996).

2.1 Interference from Polish

Consider the examples below.

(1) (a) Interference from sound	(b) Interference from spelling
E author > PE [awtor] (P *author*)	E front > PE [*front*] (E <o> > PE [o])
E balsam > PE [balsam] (P *balsam*)	E *pint* > PE [p^{j}int] (E <i> > PE [i])
E fauna > PE [fauwna] (P *fauna*)	E *mountain* > PE [mowtajn] (E <ou> > PE [ow])
E alibi > PE [a'l^{j}ibl^{j}i] (P *alibi*)	E *tomb* > PE [tomp] (E <b> > PE [b/p])

This group of errors is caused by interference from Polish pronunciation and spelling. The items in (1a) are found not only in English, but also in Polish since they are cognates, i.e. either borrowings from English or other languages (mostly Latin and Greek). Consequently, Polish learners, while speaking English, frequently pronounce them in the same way they do in their native language. A particularly numerous subgroup comprises proper nouns, such as *Presley* (pronounced as [preslej], *Streisand* ([strejsant]),*Ottawa* ([o'ttava]). The examples in (1b) are cases of spelling pronunciation in which English letters are interpreted according to the Polish letter-to-sound rules. Since Polish spelling is more phonemic than English, in the former almost all letters are pronounced. Thus, many items in this set comprise silent letters (e.g. *answer, half, tomb*).

2.2 Interference from English

Let us now turn to the errors which originate due to learners' familiarity with English.

(2) (a) Interference from sound	(b) Interference from spelling
says > PE [sejs] (*say*)	*butcher* > PE [bačer] (E <u> > PE [a], *cut*)
southern > PE ['sawdern] (*south*)	*climate* > PE [klajmejt] (E <ate> > PE [eit], *mate*)
variety > PE ['verl^{j}jeti] (*various*)	*blood* > PE [blut] E(<oo> > PE [u], *food*)
knowledge > PE ['nowl^{j}ič] (*know*)	*key* > PE [kej] (E <ey> > PE [ei], *grey*)

(2a) contains examples of errors triggered by the pronunciation of related, usually more frequently used English words. For example, the diphthong of *know* is transferred to the phonetic rendition of *knowledge*. The items in (2b) are often mispronounced due to incorrect overgeneralizations concerning English spelling-to-sound rules. For instance, the fact that <oo> in many words is pronounced as [u:], (*food, root, loose*), often leads to an erroneous assumption that this is true in

other cases as well. Consequently, *blood* is frequently mispronounced in Polish English as [blut]. It should also be added that many errors have multiple sources. For example, the word *says*, often mispronounced as [sejs], results both from the pronunciation of the infinitive with the diphthong as well as from the written form in which the digraph <ay> suggests a diphthongal realization (as in *may*).[2]

The above brief discussion on the major sources of local errors provides an explanation why such cases are generally not discussed in phonetic manuals published by big (frequently international) publishing houses. Since the majority of such books are addressed to international learners of English, their authors probably assume that local errors, resulting from the negative transfer between L1 and L2, are L1-specific, and usually focus on more general pronunciation issues of interest to a larger audience. Another reason for this neglect might be the assumption that this is a small-scale local phenomenon, unworthy of more assiduous attention. Finally, local errors seem more typical of EFL learners rather than of ESL learners as the former are exposed to written English more often than to spoken language while the opposite is true of the latter. The majority of pronunciation instruction textbooks (e.g. Celce-Murcia et al. 1996) concern primarily the ESL context.

3 A Local Phenomenon?

In order to check how widespread local errors indeed are, we have examined the way learners from several countries pronounce 50 items problematic for Poles. 15 students of 5 different nationalities, all intermediate to advanced learners of English, attending an English language school in Dublin were recorded having been asked to read a list of words prepared by the author. It has turned out that all the items were mispronounced by the majority of the subjects, regardless of their mother tongue. Table 2 presents a selection of learner versions of ten such words.[3] Of course many other phonetically incorrect renditions of the test items can be encountered.

Even this small set of data is sufficient to show that we are dealing here with a problem that concerns not only Poles, but other learners of English with different first language background as well. Interestingly, the same items turned out to be phonetically difficult for the participants, which suggests that a list of words prone to distortion by various international users of English can probably be compiled. This, however, remains a task for future research.

---

[2] Interestingly, other languages might also be a source of pronunciation errors. A case in point is the word *lieutenant*, frequently rendered by Polish learners as [lojtnant], undoubtedly due to its German version, made popular in Poland by war movies. German influences are also responsible for the mispronunciations of initial consonant clusters in names such as *Spielberg* > PE [špilberk] or *Steinbeck* > PE [štajnbek].

[3] The forms provided in the table are only rough approximations of the recorded items. This means that one symbol often stands for variety of sounds, e.g. [a] represents an open unrounded vowel whose backness varies, however, from language to language.

Table 2 Phonetic versions of ten items produced by learners of five nationalities

Tested items	Brazilian Portugese	Mauritian Creole	Mexican Spanish	Italian	Polish
meadow	[ˈmi:dow]	[‘mi:dow]/ [‘midow]	[‘mi:dow]/ [‘midow]	[‘mi:dow]	[‘ml^{j}i:dow]
climate	[kli’meit]/ [‘klimat]	[‘klaimet]	[‘klimeit]/ [‘klaimeit]	[‘klaimet]	[‘klajmejt]
guinea pig	[gi’nea pik]/ [‘gwinja pik]	[‘gaini pik]/ [gi’nea pik]	[gi’nea pik]	[gi’nea pik]	[gv^{j}i’nea p^{j}ik]
preface	[‘pri:feis]/ [‘prefa:s]/ [pre’feis]	[‘prifeis]	[[pre’feis]/ [‘pri:feis]	[‘pri:feis]	[‘pr^{j}ifejs]
colonel	[‘kolonel]	[‘kolonel]	[‘kolonel]	[‘kolonel]	[ko’lonel]
captain	[‘kaptajn]	[‘kaptejn] [‘kaptajn]	[‘kaptajn]	[‘keptajn]	[‘keptejn]
Leonard	[‘leonart]	[‘li:ona:rd] [‘leonar]	[‘li:onart]	[‘leonart]	[‘l^{j}ionart] [‘leonart]
mountains	[‘mountains]	[‘mountinz] [‘mauntin]	[‘mountains] [‘mountejns]	[‘montejns] [‘mauntejns]	[‘moutajns] [‘montejns]
vegetables	[vedʒi’tabls] [vedʒi’tebls]	[vedʒi’tejblz] [vegi’tejbl]	[‘vedʒetebls]	[‘vejdʒetebels]	[vedže’tejbls]
Turkey	[‘turkaj]	[‘tɜ:rkej]	[‘torki] [‘turki]	[‘torki]	[‘tark^{j}i]

4 The Experiment

In order to investigate the impact of local and global errors on communication via English, we have conducted a two-part experiment whose design and results are presented in the following sections.

4.1 Goals

The major goal of the experiment has been to examine English native speakers’ judgements concerning two types of mispronunciations commonly made by Polish learners: one which involves segmental and suprasegmental inaccuracies, i.e. global errors, but which is devoid of local errors, and another version which is segmentally and prosodically correct, i.e. in which there are no global errors, but where local errors can occur. Our aim was to find out how these two types of inaccuracies are evaluated by native-speakers of English in terms of comprehensibility, intelligibility, foreign accentedness and the degree of irritation/annoyance they trigger in native listeners.

A comment on the choice of English native speakers as judges is in order. In spite of numerous attempts to marginalize their role in linguistic communication via English and depriving them of the ownership of this language (e.g. Jenkins

2000; Widdowson 1994), there is no doubt that, in view of massive emigration of Poles to the British Isles that has taken place in recent years, native speakers' perception of Polish-accented English remains of vital importance in setting phonetic priorities for thousands of Polish learners.

4.2 Experimental Design: Part 1 and Part 2

The experiment consisted of two parts in which the same diagnostic passage and the same recording procedure were employed. They are summarized below.

4.2.1 Diagnostic Passage

For the purposes of the experiment a brief, five-sentence text containing 20 words commonly mispronounced by Polish learners of English was written by the author. The test items are transcribed below in their Polish English versions. In some cases several commonly occurring variants are provided. The transcription symbols refer to Polish sounds.

Disney ['dʲisnej]	*colonel* [ko'lonel]	*captain* ['keptejn/'keptajn]
worked [workt/workit]	*Turkey* [tarkʲi/'tarkej]	*radar* ['radar]
butcher ['bačer]	*nurse* [nars]	*mountains* ['mowtajns/'mawtejns]
climate ['klajmejt]	*area* [e'rʲija]	*soup* [sowp]
steak [stʲik]	*lettuce* [letʲjus]	*Cabin* ['kejbʲin]
foreign [fo'rejn]	*fruit* [fruit]	*Japan* ['džapan/'džepen]
walk [wolk]	*meadow* ['mʲidow]	

The local errors which occur in these forms stem from different sources. Most of them result from the interference of English spelling and involve pronouncing silent letters (*walk, fruit*), incorrect overgeneralization of English spelling-to-sound rules (e. <u> interpreted as P [a]) in *butcher, nurse, Turkey,* <ea> pronounced as [i] in *steak, meadow,* <ate> as [eit] in *climate,* <ai> and <ei> as [ei] in *captain, foreign, mountains*) and interference from Polish pronunciation (*radar*, *Japan*).

The diagnostic passage is presented below:

> The *Disney* family met for dinner in an old *mountain cabin* in the *area* which they liked for its *climate*. The two eldest brothers, *a colonel* and a *captain*, had just returned from *Turkey* and *Japan* where they *worked* on a *radar*. The youngest brother, who was a *butcher*, and their sister, who was a *nurse*, never went to *foreign* countries. They envied their brothers exciting jobs and asked many questions. After a meal of *soup*, *steak, lettuce* and *fruit*, they all went for a *walk* in the *meadow*.

It should be noted that the comprehension of this passage is aided by several factors. First of all, it is short and simple, without difficult vocabulary and complex

syntactic structures. Secondly, it is grammatically correct. Finally, the test items are used in sentences which provide rich contextual (semantic, lexical and syntactic) information, crucial in aiding the comprehension task.

4.2.2 The Recording

The passage presented in the previous section was recorded in two versions produced by two adult Polish men.

(a) Version A—here the speaker, an intermediate learner of English, read the passage using consonants, vowels and prosodies typical of Polish English, but without any major distortions of words provided in Sect. 4.2.1. (global errors, but no local errors).
(b) Version B—in this case the speaker, an English phonetics teacher, employed English sounds and prosodies, but mispronounced the test items in the manner indicated in Sect. 4.2.1. (local errors, but no global errors).

It should be added that both samples are characterized by a comparable speaking rate, volume and clarity of articulation, which factors, according to Derwing and Munro (1999), affect the comprehension of accented speech.

4.3 Part 1

4.3.1 Participants

In the first part of the experiment the participants were a group of 20 adult native speakers of Irish English, of mixed sex, aged between 25 and 55, all inhabitants of Dublin and having college education. The majority of them are teachers of English, employed in a language school and teaching English to foreigners from various countries. Thus, all the subjects are familiar with different versions of foreign-accented English. They all admit to having conversed in English with many Poles.

4.3.2 Evaluation

The participants were divided into two groups of ten persons each. The first group listened to Version A while the second group listened to Version B. Then they were all asked to complete a questionnaire which contained four questions. The first of them was intended to measure the comprehensibility of the recoded samples and asked how easy/difficult it was to understand the speaker. The next question concerned the evaluation of the speakers' accentedness and inquired how native/foreign the speaker's accent sounded to the subjects. Question 3 dealt with

Table 3 Experimental results—part I

Question	Speaker A (global errors, no local errors)	Speaker B (local errors, no global errors)
Comprehensibility	1.5 (very easy/rather easy to understand)	2.72 (rather difficult to understand)
Foreign accent	2.5 (slight foreign accent/rather strong foreign accent)	2.9 (rather strong foreign accent)
Annoyance	1.3 (not irritating at all/somewhat irritating)	1.9 (all options—from not irritating to very irritating)
Most important pronunciation errors	Flat intonation, no pauses in listing foods, every syllable pronounced too carefully	Mispronounced words: *soup, Turkey, nurse, worked, fruit, steak, captain, meadow,* etc.

the degree of irritation/annoyance caused by the speaker's pronunciation in the listener. Finally, the participants were requested to list those pronunciation inaccuracies which they considered particularly grave. The subjects selected one of the four answers to questions 1, 2 and 3, which were then given numerical values from 1 to 4 (the higher the score, the harsher the judgement). Subsequently mean scores for every question were calculated.

4.3.3 Results

The experimental results are summarised in Table 3.

The data show that the evaluations of the degree of comprehensibility, foreign accent and annoyance on all three counts are more severe in the case of Speaker B. Thus, Speaker A scored 1.5 points for comprehensibility with the assessment varying from 'very easy' to 'rather easy to understand,' while Speaker B's speech, with 2.72 points, was regarded as 'rather difficult to understand.' The differences in the evaluation of foreign-accentedness were somewhat less marked; Speaker A, with 2.5 points, sounded slightly foreign to half of the subjects and strongly foreign to the other half, whereas Speaker B was described by the majority as having a rather strong foreign accent, which supports observations made by other researchers (e.g. Derwing and Munro 1997) that native-speakers' accent ratings are harsher than comprehensibility ratings.

Speaker A's accent was generally viewed as 'not irritating at all' (1.3 points) while in the evaluation of Speaker B's pronunciation, with the general score of 1.9 points, there was no agreement between the subjects who employed all the options presented in the questionnaire, from 'not irritating at all' to 'very irritating.' Apparently, native listeners differed considerably in their degree of tolerance of local errors made by Speaker B, but were not annoyed by Speaker A's global errors.

It is interesting to note that in the case of Speaker A there was a strong correlation between answers to question 1 and 3. This means that the fact that Speaker A was judged as being easy to understand correlated with the opinion that his pronunciation was not annoying for the listeners, regardless of the assessment of his accentedness. As to Speaker B, here a correlation was noted between answers to questions 1 and 2. In other words, the ease or difficulty of understanding this speaker correlated with the assessed degree of his foreign accent.

Finally, the last question was open and asked the participants to list those pronunciation errors in the two samples which they considered particularly grave. The majority of comments on Speaker A's version concerned prosodic issues. He was claimed to have flat intonation, to make no pauses were they were required and to use too careful pronunciation ("e.g. each and every syllable pronounced too carefully and in a precise manner, there is no flexibility that a more fluent and confident speaker would have"). Interestingly, no segmental inaccuracies of this speaker were pointed out ("e.g. I did not really find too many errors in his pronunciation"). In the case of Speaker B the participants presented lists of mispronounced words with frequent comments, such as "I would normally have no problem understanding a Polish accent, but this one was difficult."

4.4 Part 2

4.4.1 Participants

Twenty subjects participated in the second part of the experiment. They were all educated (university degree) adult (aged 28–57) native speakers of English English, of mixed sex, living in London and working at an exhibition centre. They admitted having minimal contacts with Poles and Polish-accented English.

4.4.2 Evaluation

The participants formed two groups of ten persons each whose task was to transcribe orthographically the diagnostic passage in Speaker A's version (Group 1) and Speaker B's version (Group 2) in order to examine how intelligible the two versions were. Breaks were made between sentences to give the subjects sufficient time for writing them down.

At this point it should be added that while some researchers use the terms *comprehensibility* and *intelligibility* interchangeably, others make a distinction between them. According to Field (2005, 400), "intelligibility is measured by the ability of judges to transcribe the actual words of an utterance, comprehensibility by an overall rating of how easy it is to understand a given speaker." These two concepts, together with accentedness, are defined by Derwing and Munro (2005, 385) in Table 4.

Table 4 Intelligibility, comprehensibility and accentedness (Derwing and Munro 2005, 385)

Term	Definition	Measure
Intelligibility	The extent to which a listener actually understands an utterance	Transcription task (% words correct)
Comprehensibility	A listener's perception of how difficult it is to understand an utterance	Scalar judgement task (from extremely easy to understand to extremely difficult to understand
Accentedness	A listener's perception of how different a speaker's accent is from that of the L1 community	Scalar judgement task (from no accent to extremely strong accent)

Comprehensibility is thus a subjective evaluative judgement, measured best by means of a rating scale whereas intelligibility is a more objective recognition of words and utterances, which can be tested by rewriting, rephrasing or gap filling. Researchers emphasize that comprehensibility does not always go hand in hand with intelligibility, since in listening comprehension tasks it is usually not necessary for listeners to understand every single word to pick up the gist of the message. Therefore, these two skills should be measured separately. The second part of the experiment described here is an attempt to examine the intelligibility of the two samples of Polish-accented speech to native speakers of English.

4.4.3 Results

The dictation of Speaker A's passage resulted in completely accurate transcriptions written by all the participants. This means that in this case the intelligibility score was 100 %.

The transcription task of Speaker B's version turned out to be considerably more difficult with the mean intelligibility score of 76 %. In each transcription the percentage of words matching exactly those in the diagnostic passage was calculated. Next the mean listener scores were averaged. The words and phrases whose understanding was particularly problematic are presented in Table 5 in the order of decreasing difficulty. The provided figures refer to the percentage of mistranscribed items.

These results demonstrate that the two experimental recordings differ strikingly in terms of their intelligibility, with Speaker A's version being fully intelligible and Speaker B's version remaining in many cases unintelligible to native English listeners.

It is interesting to compare the intelligibility scores of the test items, which were calculated by counting the percentage of forms exactly matching those found in the original versions. As is well-known, apart from the phonetic and phonological features of a given word, other factors relevant for intelligibility include, among other things, contextual transparency as well as the listener's syntactic and lexical knowledge. The least intelligible item was *Disney*, mistranscribed by all the participants, which can be accounted for by the phonetic unpredictability of

Table 5 Experimental results—part II

Test item	Subjects' versions
the Disney family (100 % incorrect)	Dista, Distay, Distin, Destain, Distone
they worked on a radar (90 %)	they walked on a rudder, they walked on a ladder, they walked on water
sister was a nurse (80 %)	was an arse, was an ass, was an aunt
a meal of soup, steak (70 %)	soap stick, steep, a reel of soap stick
Turkey (60 %)	talky, Takhi, Taki
colonel (50 %)	a colonial, Appolonia
captain (50 %)	katung
butcher (50 %)	bachelor
fruit [fruit] (50 %)	Freud
walk in the meadow (40 %)	walk in the middle
Japan (40 %)	Chatham, Chapan
for its climate (30 %)	for their environment, for its time it, for its timing
lettuce (30 %)	the juice, let choose
cabin (30 %)	cave
in the area they liked (20 %)	is there a rear, the rear the light
foreign (10 %)	(?)
mountain (0 %)	(?)

surnames and other proper names in general. This was additionally reinforced by the rarity of morpheme internal occurrences of clusters of voiceless fricatives and nasals present in the Polish English version [d'isnej] and perceived by all the listeners as the [st] sequence, far more frequent in this position in English.

The phrase *they worked on a radar*, with 90 % of incorrect answers, came second in terms of difficulty to understand, due to mispronunciation of both the verb and the noun, followed by *sister was a nurse* (80 %) and *a meal of soup, steak* (70 %). The easiest words to decipher turned out to be *foreign, mountain* and *area* (0–20 % of incorrect responses). This result means that not all types of local errors represent the same level of processing difficulty for the listeners, which is strictly connected with their phonological make-up (e.g. the word *mountain*, even when grossly distorted, cannot be easily confused with any other item) and contextual transparency. Thus, the expression *foreign countries*, for instance, created a lexical, syntactic and semantic context which made the adjective comprehensible in spite of its deviant phonetic realization. Clearly, further tests with these and other expressions used in different contexts are needed to establish their intelligibility values.

It is also worth pointing out that, when faced with strange-sounding and unintelligible expressions, listeners adopted several different decoding strategies. The first and the simplest of them consisted in placing a question mark or a hyphen in place of a problematic item, which means the lack of any attempt to interpret it. Another strategy involved writing down exactly what the listener heard, even if it meant creating an utterance which was either ungrammatical (e.g. *for its climate* rendered as *for its time it*), semantically uninterpretable (*lettuce* transcribed as *let choose*) or both (*in the area they liked written down as is there a rear* or as the rear the light). Finally, some of the

listeners tried to make sense of a given expression by changing it completely to fit the context (*a mountain cabin* interpreted as *a mountain cave*).

5 Conclusion

This paper has argued that the major obstacle to successful communication between foreign learners and native-speakers of English is posed not so much by global segmental and prosodic inaccuracies as by idiosyncratically deviant words, i.e. local errors, which abound in foreign-accented English. Such items, generally disregarded in intelligibility research and in pronunciation training materials, have been shown to affect negatively native judges' ratings of Polish-accented English in terms of its comprehensibility, intelligibility, foreign-accentedness and the degree of irritation triggered in the listeners far more than corresponding ratings of global errors.

Pedagogical implications of this research are clear; it is not the articulation of individual sounds and prosodic patterns of English that should be the focus of communicatively-oriented pronunciation instruction, but rather whole words prone to be commonly mispronounced by language learners[4].

Acknowledgments This paper could be completed due to the help received from Graz Kalenik and Dariusz Bukowski, to whom I am deeply grateful.

References

Abercrombie, David.1949. Teaching pronunciation. English Language Teaching 3, 113–122.

Avery, Peter. and Susan Ehrlich. 1992. *Teaching American English Pronunciation.* Oxford: Oxford University Press.

Celce-Murcia, Marianne, Donna M. Brinton and Janet M. Goodwin. 2010. *Teaching Pronunciation. A Reference for Teachers of English to Speakers of Other Languages.* Cambridge: Cambridge University Press.

Cruttenden, Alan. 1994. *Gimson's Pronunciation of English.* London & New York: Arnold.

Derwing, Tracey. and Murray Munro. 1997. Accent, intelligibility and comprehensibility. Studies in Second Language Acquisition *20, 1–16.*

Derwing, Tracey and Murray Munro. 2005. Second language accent and pronunciation teaching: a research-based approach. TESOL Quarterly *39, 3: 379–397.*

Field, John 2005. Intelligibility and the listener: the role of lexical stress. TESOL Quarterly *39, 3: 399–423.*

Jenkins, Jennifer. 2000. *The Phonology of English as an International Language.* Oxford: Oxford University Press.

Kennworthy, Joanne. 1987. *Teaching English Pronunciation.* London: Longman.

Munro, Murray and Tracey Derwing. 1999. Foreign accent, comprehensibility and intelligibility in the speech of second language learners. Language Learning *49, 1: 285–310.*

[4] Some pedagogical suggestions as to how local can be erradicated are found in Szpyra-Kozlowska and Stasiak (2010).

Sobkowiak, Włodzimierz. 1996. *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.
Szpyra-Kozłowska, Jolanta. 2005. Lingua Franca Core, phonetic universals and the Polish context. In *English Pronunciation Models: A Changing Scene,* eds K. Dziubalska-Kołaczyk and J. Przedlacka, 151-176, Bern: Peter Lang.
Szpyra-Kozłowska, Jolanta & Sławomir Stasiak. 2010. From focus on sounds to focus on words in English pronunciation instruction. *Research in Language* 8. 163–177.
Widdowson, Henry. 1994. The ownership of English. TESOL Quarterly *28/2, 377–389.*

Foreign Accent Ratings in Third Language Acquisition: The Case of L3 French

Magdalena Wrembel

Abstract The present paper constitutes a part of a larger-scale project aimed at investigating the sources and directions of cross-linguistic influence in the phonological acquisition of a third language. To this end, foreign accent ratings were performed on a pool of speech samples of L3 French recorded by L1 Polish subjects, proficient in L2 English, with different competence levels in French. The perceptual judgements of L3 French, performed online by expert raters, involved the ratings of accent, intelligibility, acceptability and confidence level as well as the identification of the subjects' L1. The study revealed interesting patterns of correspondence between different rating parameters, and demonstrated a high level of inter-rater consistency. The findings pointed to the multilingual subjects' first language as the prevailing source of cross-linguistic influence in the phonological acquisition of an L3, however, the L2 influence was also found to be considerable, judging on the basis of the raters' identification of the subjects' L1. The results appear to be consistent with the assumption of a combined cross-linguistic influence (cf. De Angelis, Third or additional language acquisition. Multilingual matters, Clevedon, 2007) thus substantiating the existence of both native and non-native interference in L3 phonology, although its strength is hypothesised to be dependent on the typological relatedness of the respective language combinations.

1 Introduction

Until fairly recently, empirical investigations into the phonological acquisition of a foreign language have focused primarily on the influence of the native tongue on the acquisition of a second language. But for the past few years, there has been a

M. Wrembel (✉)
Adam Mickiewicz University, Poznań, Poland
e-mail: magdala@amu.edu.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_3,

growing recognition that Third Language Acquisition (TLA) is a subfield of inquiry separate from Second Language Acquisition (SLA) and scholars have started to differentiate between the acquisition of an L2 and other subsequent languages (cf. Cenoz et al. 2001; Cenoz and Jessner 2000; De Angelis 2007). The phonology of third language (L3) acquisition constitutes a particularly young discipline and research in this area has been limited in scope compared to that on L3 lexis and morphosyntax (cf. Cabrelli Amaro and Rothman 2010).

The complexity of cross-linguistic influence (CLI) has become one of the most frequently studied aspects of TLA, with investigations into the nature of L3 interphonology being no exception. Two opposite perspectives on the potential sources of CLI in L3 have been endorsed; one pointing to the native L1 phonology as having the greatest impact on third language acquisition (Ringbom 1987; Llisterri and Poch-Olivé 1987; Pyun 2005), and the other recognizing the non-native language, i.e. the earlier acquired L2 as the prevailing source of interference in L3 production, thus leading the so called 'foreign language effect' or L2 status (Hammarberg and Hammarberg 1993, 2005; Llama et al. 2010; Wrembel 2010).

A relatively limited number of studies investigating this phenomenon has been conducted to date, and their results appear to conflict. The traditionally acknowledged claim that foreign language acquisition, especially in the area of L3 phonology, is subject to a strong interference from the first language was attested by Ringbom (1987), Llisterri and Poch-Olivé (1987) and Pyun (2005). According to Ringbom, all L3 learners tend to retain an L1-based accent, particularly in the realm of intonation, irrespective of their L3 proficiency levels. He further claims that L2 transfer in L3 interphonology is infrequent and may be governed by the recency and intensity of L2 use. Llisterri and Poch-Olivé (1987) demonstrated experimentally that in the case of Catalan/Castilian bilinguals learning a third language there is no interference from their L2 in the production of vowels and fricative consonants in L3 French compared to the monolingual controls. The observed transfer in L3 was explained by the authors by referring to the acoustic features of the subjects' L1 sound system. Negative transfer from the L1 phonology was also evidenced as the strongest source of CLI in third language acquisition in a study by Pyun (2005) on the multilingual speakers with Korean L1, English L2 and Swedish L3. However, he identified also other sources of phonological knowledge in the subjects' interlanguage, involving an interplay of various L1, L2 and L3 rules.

On the other hand, several studies to date testify to the complex interference between foreign languages acquired as second and third and point to the existence of the so called foreign language effect that tends to outweigh L1 transfer in L3 phonological acquisition. The earliest reported investigations into the sources of CLI in L3 involved impressionistic case studies (Chamot 1973) or diary studies of a multilingual informant (Rivers 1979), both of which demonstrated non-native phonetic interference in the L3 performance. With the advent of studies using a more controlled experimental design, further evidence was provided for the complex nature of L3 interphonology, which is understood as the evolving system of phonological rules that is restructured during the successive stages of foreign language acquisition (cf. Carlisle 1991, 76). A longitudinal study of a single

multilingual speaker with L1 English, L2 German and L3 Swedish (Hammarberg 2001 is a case in point (Hammarberg and Hammarberg 1993, 2005).The study was based on accent judgements by 3 native raters on the informant's samples of L3 Swedish as well as an auditory and acoustic analysis of these samples performed by the authors at different stages of interlanguage development. The findings pointed to a prevalent interference of the non-native language at an early stage of L3 acquisition (when the informant's performance in L3 Swedish was strongly L2-accented) as opposed to more native interference at the later stage. The results of accent judgements were corroborated by the auditory analyses by the experimenters of the informant's performance in L3 Swedish. Wrembel (2010) confirmed these findings in a perceptual accent judgement study of foreign accentedness of L3 learners of English with L1 Polish and L2 German. The results reflected a significant tendency to identify the subjects as L2-accented at the initial stages of L3 learning, which diminished with the advancement of L3 proficiency. The author claims that results reflect a cumulative effect of several variables including the foreign language effect, typology (the subjects' L3 English and L2 German both being Germanic languages), and the recency and intensity of use of L2 German. Similar trends were attested in a study on VOT acquisition in L3 by Llama et al. (2010), who concluded that it is the L2 foreign language effect (or the so called 'L2 status') rather than typological relatedness that constitutes the main factor determining the source of transfer of VOT patterns for L3 Spanish learners in L1 French/L2 English and L1 English/L2 French language combinations.

Investigations into the possible sources of CLI in L3 have also yielded inconclusive results in some of the most recent studies in the field. Gut (2010) found conflicting evidence for L2-L3 cross-linguistic influence as her data suggest that while L2 and L3 proficiency and language distance might be conditioning factors for L2–L3 influence, the effects of the phonological properties of the L3 are stronger than cross-linguistic influence. Wunder's (2010) results were also mixed, pointing to some non-native cross-linguistic influence on L3 phonology and combined L1 German and L2 English influence on L3 Spanish aspiration patterns. Finally, Wrembel's (2012) foreign accent study involved a mirror design with respect to the language combinations employed in the present study which was achieved by exchanging the L2/L3 language combinations, i.e. from L1 Polish/L2 French/L3 English in the present study to L1 Polish/L2/English/L3 French in the other study. Wrembel (2012) demonstrated that L1 transfer from the subjects' Polish phonology dominated in their performance in L3 English irrespective of the stage of language proficiency in L3. However, a partial reliance on L2 French phonetic encoding was also reported.

2 Experiment

The present experiment is part of a larger study aimed at investigating the sources and directions of cross-linguistic influence in L3 phonological acquisition in various language combinations. The current study focuses on the acquisition of L3

French by subjects with L1 Polish and L2 English, and it follows a mirror design of a previous foreign accent study by the same author which involved L1 Polish, L2 French and L3 English (Wrembel under review).

Foreign accent ratings have been widely applied in studies on second language acquisition, e.g. Højen (2000), Piske et al. (2001), Gallardo del Puerto et al. (2007). However, to the best of my knowledge, they have not been applied extensively in research on third language phonological acquisition with the exception of Hammarberg and Hammarberg's case study (cited above), in which L3 accent ratings were performed by three raters at two different stages of interlanguage development.

The technique of foreign accent rating was selected in order to generate accent ratings in a third language as well as to explore the L3 speakers' degree of intelligibility and acceptability, and the raters' level of confidence. Moreover, it was expected to contribute to the investigation of potential sources of cross linguistic influence in TLA as the raters were to identify the multilingual subjects' first language based on the perceptual evaluation of their L3 performance, and to point to the phonetic/phonological features that led to the perception of a foreign accent in particular L3 speakers.

2.1 Research Design

Twenty speech samples of L3 French (30–40 s each) were randomly selected from a small corpus of non-native French recorded by the author and were subject to foreign accent ratings performed by 15 expert judges. The raters included 10 native speakers of French as well as five non-native speakers with a near-native proficiency level in French, all of whom were Polish university lecturers. The judges' mean length of French teaching experience was 12 years and they all had previous phonetic training, irrespective of their native versus non-native status. The stimuli were recorded using CoolEdit96 as 16-bit mono files at 16000 Hz sampling frequency. The subjects were recorded performing language tasks in two conditions: a read-on-your-own task and a spontaneous speech.

The experiment was designed on the e-learning Moodle platform so that it could be performed by the raters via the Internet at their convenience. The raters were assigned an access code and after the completion of the online questionnaire, it was submitted automatically to the present researcher. The online questionnaire was divided into two parts, the first one elicited the raters' personal information, including their native versus non-native status, their teaching experience in years, and their knowledge of foreign languages.

The second part of the experiment consisted of 20 identical subsections, each equipped with a recording (i.e. ten recordings of read texts and ten recordings of spoken texts) and a set of six questions. The judges were asked to perform the following tasks:

1. rate the samples for an overall degree of foreign accent on a 6-point scale (1 = strongly accented, 6 = native-like accent),

2. evaluate how intelligible the speaker sounds (on a 6-point scale from 1 = unintelligible, 6 = totally intelligible),
3. assess the degree of confidence in their rating judgement (on a 6-point scale from 1 = uncertain, 6 = certain),
4. rate how irritating/acceptable the speaker sounds (on a 6-point scale from 1 = very irritating, 6 = totally acceptable),
5. identify the speakers' L1 (i.e. select a language from an open list presented in the rating questionnaire),
6. point out the phonetic/phonological features that contribute to the perceptual impression of foreign accent in particular speakers.

2.2 Participants

The recorded subjects included 10 native speakers of Polish with a very good command of English as their L2 and with rather limited competence in French as their L3. One native speaker of French with L2 English was included as a control. The Polish participants were students of the School of English at Adam Mickiewicz University in Poznań. Their proficiency level in French, assessed on the basis of competence level tests administered at the university, ranged from elementary (A2 Basic Speakers, according to CEFR) to intermediate (B1/B2 Independent Speakers), whereas their competence in L2 English was more advanced as they were all at the C1 level (i.e. Proficient User).

A language learning biography of each multilingual speaker was constructed in order to control better for various factors contributing to the complex nature of cross linguistic influence. The relevant data were collected by means of a questionnaire administered individually after the recording session. The questions covered the following: (1) age of acquisition of each non-native language, (2) proficiency level in all non-native languages, (3) sequence of acquisition of all languages, (4) amount of formal instruction in non-native languages (in years and hours per week), (5) natural exposure to non-native languages (e.g. stays abroad), (6) number of languages known to the speaker, (7) metalinguistic awareness of positive or negative transfer from non-native languages in L3 oral production. The analysis of language biographies resulted in a fairly homogenous profile with respect to the participants' L2 English, however, they differed significantly in proficiency level in L3 French.

2.3 Hypotheses

The following hypotheses were put forward in the present study:

H1: Accent judgements are correlated with speaker proficiency level in L3.

H2: There are patterns of correspondence between the rating scores on accent, intelligibility, acceptability and raters' confidence.

H3: The subjects are identified as native speakers of Polish or English on the basis of their L3 performance depending on their proficiency level in L3 French.

H4: There are differences in accent judgements due to the task and inter-rater variability.

3 Results

3.1 Foreign Accent Ratings

Perceptual judgments of foreign accent in L3 French involved several rating parameters, namely:

1. degree of perceived foreign accent,
2. individual speaker intelligibility,
3. confidence level of judges,
4. degree of irritation or acceptability evoked by the speakers' accents.

Table 1 presents the results for all the rating parameters including total mean values as well as individual values for 20 speech samples. The highlighted squares indicate values significantly higher than the remaining ratings based on the comparison of means which was calculated with the use of a two tailed *t* test with Bonferroni's correction ($p < 0.05$).

As far as foreign accentedness is concerned, the mean total rating of the L3 French samples was 3.3 on a 6-point scale (1 = strongly accented, 6 = native-like accent). Participants whose results were significantly higher than the remaining ratings were speakers 4, 7, 8, 9, 13, 17, 19 and 20. All of these participants were at the intermediate level. Speaker 8, who received the maximum ratings, was the native French control.

The second perceptual judgment task related to the speakers' intelligibility and the mean rating was 4.5 on a 6-point scale (ranging from 1 = unintelligible, 6 = totally intelligible). The intelligibility ratings were on average higher than accent ratings. Statistical differences based on the comparison of means were found for intermediate level speakers (i.e. speakers 3, 4, 7, 9, 11, 13, 17, 19, 20) and for the native control (speaker 8).

The level of judges' confidence when performing the accent rating tasks had a mean value of 5.4 on a 6-point scale (from 1 = uncertain, 6 = certain) and remained fairly stable across the speakers as no statistically significant differences were found in the certainty ratings with respect to different speakers.

As far as the acceptability rating is concerned, the mean score for L3 French samples totalled 4.2 on a 6-point scale (1 = very irritating, 6 = totally acceptable). The samples which scored significantly higher in terms of acceptability included intermediate level speakers (i.e. speakers 3, 4, 7, 9, 11, 13, 17, 19, 20) and the native control (speaker 8), as in the accent and intelligibility ratings.

Table 1 Mean accent scores for four rating tasks for L3 French on a 6-point scale

Speaker	Task			
	Foreign accent	Intelligibility	Certainty	Acceptability
M	3.3	4.5	5.4	4.2
1	1.4	2.9	5.7	2.7
2	1.4	2.3	5.5	2.3
3	3.5	5.3	5.3	5.1
4	4.3	5.3	5.4	5.1
5	3.3	4.4	5.5	3.7
6	1.5	2.8	5.5	2.5
7	4.3	5.4	5.5	5.5
8	6.0	6.0	4.3	6.0
9	4.1	5.1	5.1	5.1
10	1.6	2.7	5.7	2.7
11	3.7	4.7	5.1	4.1
12	2.5	4.4	5.2	3.9
13	4.5	5.7	5.5	5.5
14	2.2	4.2	5.7	3.4
15	2.9	4.1	5.2	3.6
16	2.6	4.2	5.6	3.9
17	4.3	5.7	5.2	5.4
18	2.3	4.2	5.5	3.1
19	4.7	5.8	5.1	5.5
20	4.1	5.3	5.6	5.1

The highlighted squares indicate significantly different scores at the 0.05 level

3.2 L1 Identification

The subsequent rating task required the judges to identify the speakers' native language (L1) on the basis of their L3 French performance. To this end, the judges were to select one from an open list of L1 labels presented in the online questionnaire, including English, French, German, Polish, Spanish and other.

The overall results presented in Table 2 indicate that the subjects were correctly identified as native speakers of their respective languages, i.e. L1 Polish or L1 French (in case of the control subject) on average in 50 % of the cases. They were identified as English on average in 29 % of the cases. The control native speaker of French was identified correctly in 100 % of the cases. Statistically significant differences in correct L1 identification based on the comparison of means (i.e. a two tailed *t* test) were revealed for speakers 1, 5, 8, and 14 for whom the values for correct L1 identification ranged between 70 and 100 %. The subjects with significantly strong identification as L1 English included Speakers 6 and 20 with the respective scores of 73 and 87 %.

After the exclusion of the scores for the French control, the identification percentages were calculated for the Polish subjects only. Based on their L3 French performance, 47 % were identified correctly as having L1 Polish; 30 % as having

Table 2 Identification of speakers' L1 (in %) based on the performance in L3 French (correct identification as Polish vs. identification as english)

Speaker	L1 identification	
	Correct identification (%)	Identification as English (%)
M	50	29
1	100	0
2	53	0
3	40	40
4	40	27
5	80	13
6	20	73
7	33	27
8	100	0
9	47	40
10	33	33
11	47	33
12	47	20
13	33	33
14	73	27
15	67	20
16	33	27
17	40	7
18	53	33
19	47	40
20	7	87

NB. Identification as 'other L1' not included in the analysis

English as their L1, and in 23 % of the cases a different native tongue was identified. The identification of other L1s is excluded from the present analysis as the results were very scattered and showed no regularities.

3.3 Speakers Variables

3.3.1 Accent Ratings Versus Proficiency Level and Performance Mode

The results of accent judgements were further analysed with respect to speaker characteristics, i.e. their proficiency level in L3 French (i.e. elementary vs. intermediate) and the recorded performance mode (read vs. spoken text). The mean scores are presented in Table 3.

The comparison of means calculated on the basis of a two tailed t test point to statistically significant differences ($p < 0.05$) for three rating parameters in L3 French (i.e. accentedness, intelligibility and irritability) with regard to both the subjects' proficiency level in French and the performance mode.

The intermediate subjects outperformed those at an elementary level with respect to proficiency in L3 French in all the rating parameters apart from the

Table 3 Mean ratings with respect to the subjects' proficiency in L3 French and the performance mode

	Total		Level		Mode	
	M	SD	Element	Intermed	Read	Speak
Accent	3.3	1.3	2.5	3.9[a]	2.9	3.3[a]
Intelligible	4.5	1.1	3.9	5.2	4.1	4.8[a]
Certain	5.4	0.3	5.4	5.4	5.4	5.4
Irritating	4.2	1.2	3.5	4.9[a]	3.9	4.4[a]

Highlighted squares point to statistically significant differences
[a] $p < 0.05$

judges' confidence level which remained stable for both proficiency levels ($M = 5.4$). Mean accent ratings were significantly higher for the more advanced subjects ($M = 3.9$ vs. $M = 2.5$) rendering the results consistent with the participants' proficiency levels. The intelligibility ratings followed the same pattern with intermediate subjects scoring on average higher than lower proficiency subjects ($M = 5.2$ vs. $M = 3.9$). In terms of tendency to irritate, the more proficient speakers were found to be more acceptable than their elementary counterparts ($M = 4.9$ vs. $M = 3.5$).

Similar patterns were observed in the analysis of the performance modes (i.e. read vs. spoken text). On the whole, spoken samples received significantly higher ratings than the read text in case of three parameters: i.e. accentedness ($M = 3.3$ vs. $M = 2.9$); intelligibility ($M = 4.8$ vs. $M = 4.1$); and irritability ($M = 4.4$ vs. $M = 3.9$). No significant differences were observed with regard to the certainty level of the raters, which totalled $M = 5.4$ irrespective of the speakers' performance style.

3.3.2 L1 Identification Versus Proficiency Level and Performance Mode

The results of identification of the subjects' L1 on the basis of their L3 performance were further analysed with respect to their proficiency level (elementary vs. intermediate) and the performance mode (reading vs. speaking) (see Table 4).

In order to investigate the statistical significance of differences between the mean results of the identification as L1 Polish (i.e. correct identification) versus the identification as L1 English (English being the L2) at different proficiency levels in L3 French and in different performance modes, a Pearson Chi-square test was calculated (see Table 5). The test showed a significant difference in the identification as L1 Polish at different proficiency levels ($p < 0.01$) with the elementary subjects being more frequently correctly identified as Polish native speakers (53 %) compared to a significantly lower percentage of identifications as L1 English (28 %). At the intermediate level, however, the difference between identification as Polish versus English was not significant (38 vs. 34 %). No statistically significant differences were found in case of the performance style and identification type, although the read samples tended to be more often

Table 4 L1 identification versus proficiency level and performance mode

	Total (%)	Level		Mode	
		Element (%)	Intermed (%)	Read (%)	Speak (%)
Identification as Polish	47	53	38	49	44
Identification as English	30	28	34	28	32

Table 5 Pearson's Chi-square test

		Level	Mode
Identification as Polish	$\chi 2$	6.27	3.52
	df	2	2
	sig	0.04[a]	0.17
Identification as English	$\chi 2$	1.55	3.98
	df	2	2
	sig	0.46	0.13

[a] $p < 0.05$ for L1 identification versus proficiency level and performance mode

identified correctly as L1 Polish (49 %) than the spoken samples (44 %), and the reverse was true for the identification as English (32 %—spoken vs. 28 %—read samples).

The significant difference in L1 identification for subjects on different proficiency levels in L3 French is presented visually in Fig. 1.

3.4 Inter-Rater Variability

A further analysis was aimed at inter-rater variability to investigate a potential correlation between the rating scores on the four parameters (i.e. foreign accent, intelligibility, certainty and acceptability) and the raters' variables, including: (1) native versus non-native status, (2) phonetic training, and (3) the length of the teaching experience.

Table 6 presents mean ratings for the first two variables. Unlike in the author's previous studies, in case of the present study all the judges had phonetic training, therefore no variety was demonstrated in this respect. The nativeness variable was shown to have affected the ratings as native raters tended to assign lower score for accent, intelligibility and irritation compared to non-native judges, however, this difference proved statistically significant only for intelligibility ($M = 4.3$ vs. $M = 4.9$). The comparison of means calculated on the basis of a two tailed t test pointed also to statistically significant differences ($p < 0.05$) in the level of certainty with native judges being more certain of their ratings ($M = 5.6$ vs. $M = 4.9$).

To analyse the impact of the third variable, i.e. the length of French teaching experience on the ratings, a Spearman's correlation was performed (see Table 7).

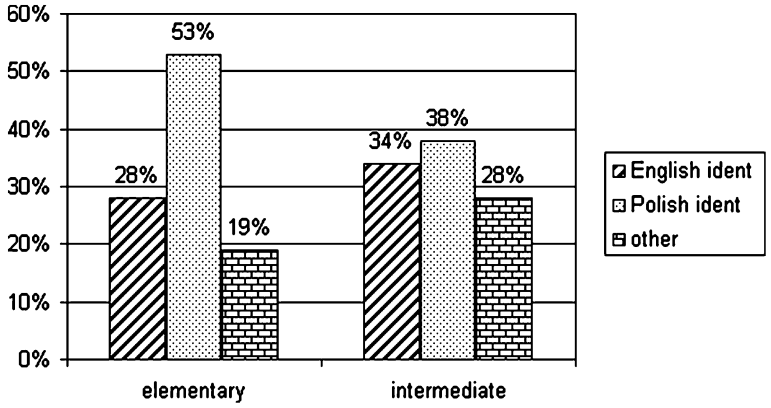


Fig. 1 L1 identification in L3 French for different proficiency levels

Table 6 Mean ratings with respect to raters' variables

Raters' variables	Native status		Phonetic training
	Non-native	Native	
Accent	3.5	3.1	3.3
Intelligible	4.9[a]	4.3	4.5
Certain	4.9	5.6[a]	5.4
Irritating	4.3	4.1	4.2

[a] $p < 0.05$

The results pointed to a weak inverse correlation between the raters' length of teaching experience and their level of confidence, i.e. the more experienced judges assigned themselves lower points for the rating certainty. No correlations were observed between the raters' French teaching experience and their ratings on accent, intelligibility and acceptance.

3.5 Rating Consistency

In order to investigate the rating consistency separate tests for inter-rater and inter-speaker reliability were administered. Cronbach's alpha was calculated to test the inter-rater reliability which was found to be high for all the four rating parameters (foreign accent 0.89; intelligibility 0.92; confidence level 0.94; acceptability 0.89). In case of L1 identification of the subjects, the inter-rater reliability ranged from moderate for the correct identification as Polish (Cronbach's alpha 0. 49) to low for identification as English (0.24).

A Spearman's correlation analysis was performed to test the inter-speaker reliability, i.e. the consistency of ratings assigned to the same subjects for their

Table 7 Spearman's correlation between the raters' years of teaching experience and their respective ratings

			Accent	Intelligible	Certain	Irritating
rho Spearman *r*	Years of teaching	Correlation coefficient	−0.005	0.096	−0.18[a]	0.051
		Sig.	0.93	0.09	0.002	0.38
		N	300	300	300	300

[a] $p < 0.01$

separate read and spoken performances (e.g. Accent vs. Accent2). Table 8 presents significant correlations on the 'match-guise' ratings performed by the judges with respect to the four rating parameters. The strongest correlations were found for the acceptability (0.71) and confidence level (0.68), however, the remaining parameters also demonstrated moderate correlations, i.e. intelligibility (0.6) and the degree of foreign accent (0.58).

3.6 Foreign Accent Features

The raters were also asked to identify phonetic/phonological features that contribute to the perceptual impression of a foreign accent of particular subjects. The most frequently identified features of a foreign accent in L3 French included vowels (24 %), consonants (10.8 %), intonation (9.5 %), rhythm (8.5 %), problems with the French [R] (8 %), grapho-phonemic correspondence (7.5 %) and nasality (7.3 %) (see Fig. 2).

4 Discussion

The first hypothesis put forward in the present study assumed that accent ratings are correlated with the subjects' proficiency levels. The results corroborated this claim since the ratings corresponded to the subjects' proficiency levels and individual foreign accent scores assigned by the judges reflected the subjects' respective command of L3 French. The mean foreign accent ratings for the elementary level subjects totalled $M = 2.5$ (on a 6-point scale), whereas the accent ratings for the intermediate learners were significantly higher $M = 3.9$. The native French control (Subject 8) received maximum scores in accent, intelligibility and acceptability. On the whole, the findings lent validity to the performed foreign accent ratings as they reflected the subjects' actual proficiency levels in L3 French.

The second hypothesis assumed the existence of patterns of correspondence between the rating scores on accent, intelligibility, acceptability and rater confidence. The lowest mean scores were assigned for foreign accent ($M = 3.3$), with all of the intermediate level subjects scoring significantly higher than the elementary ones. The acceptability scores were on average higher ($M = 4.2$) and they

Table 8 Spearman's correlation between the subjects' two separate recordings and judges' respective ratings

			Accent_2	Intelligible_2	Certain_2	Irritating_2
rho Spearman (*r*)	Accent/ Intelligible/ Certain/ Irritating	Correlation coefficient	0.58[a]	0.6[a]	0.68[a]	0.71[a]
		Sig.	0.00	0.00	0.00	0.00
		N	135	135	135	135

[a] $p < 0.01$

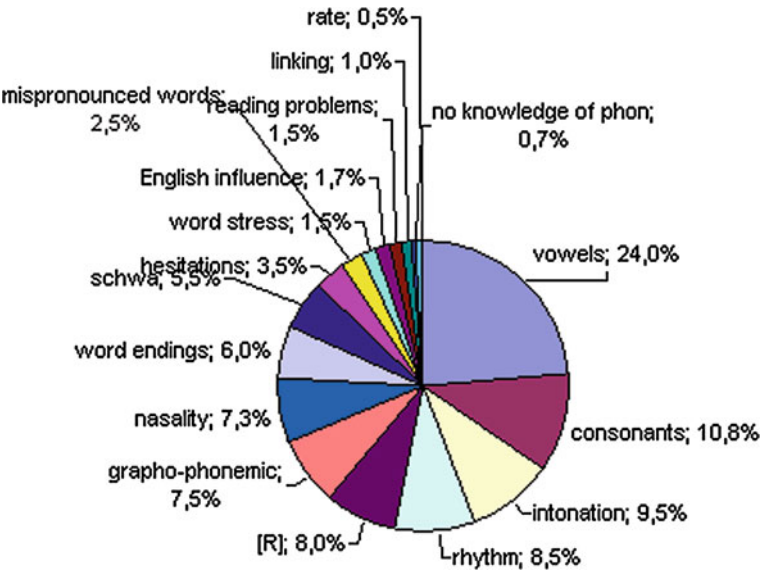


Fig. 2 Foreign accent features in L3 French according to the raters

corresponded closely to the accent scores, i.e. the same intermediate subjects were rated as more acceptable as those who received higher accent ratings. The mean intelligibility scores were the highest of the three rating parameters ($M = 4.5$) and they corresponded closely to the scores on accent and acceptability with all the intermediate level subjects outperforming their elementary counterparts. The level of rating confidence remained essentially stable ($M = 5.4$) and did not show any level-dependent variation.

The hypothesis with the greatest bearing for the present study was that the Polish subjects performing in L3 French will be identified by the raters as having either Polish or English as their L1 due to the high proficiency in L2 English. The present results show that the native language appeared to be the strongest source of cross-linguistic influence in the acquisition of L3 phonology as in 47 % of the cases the subjects were identified correctly as having Polish as their L1. Nonetheless, the non-native language also seems to have exerted a considerable influence on the L3 interphonology since the percentage of identifications as

English was high (30 %), presumably due to the subjects' advanced proficiency in this language.

The emerging patterns of L1 identification with respect to the subjects' proficiency level were quite different from the ones reported in the literature. Contrary to results from most studies cited above, the percentage of correct identification as Polish was significantly higher at the L3 French elementary level (53 %) than at the intermediate level (38 %). Consequently, the identification as English, i.e. the perception of L2-accented speech was more frequent at the higher than at the lower proficiency level (34 vs. 28 %), though the difference was not statistically significant. The identification of other languages as potential L1s amounted to 28 % for the intermediate and 19 % for the elementary L3 learners. The results, however, were very diverse and no clear regularities in other L1 identifications were observed.

The effect of the performance mode in L3 French was not found to be statistically significant for the identification of the subjects' L1 as was expected, based on the author's previous findings in which spoken L3 samples exhibited more L2-accented features whereas read L3 samples were more frequently judged as being L1-accented (Wrembel 2010). However, the tendencies observed in the present study were in accordance with those reported earlier as they evidenced a higher percentage of English identification in the speaking mode versus a higher percentage of Polish identification in the reading mode.

The findings neither support nor disconfirm Hammarberg and Hammarberg's claim that it is the other non-native language (L2) rather than the mother tongue that constitutes the stronger source of transfer at the initial stages of L3 phonological acquisition thus resulting in L2-accented speech in L3 performance. The present results based on foreign accent ratings lead to the conclusion that it is the native language that is prevalent as the source of cross-linguistic influence in the phonological acquisition of L3, however, the influence of the non-native tongue seems to be also considerable. The strength of the L2 effect factor was greater than in the mirror design study (Wrembel 2012), however, Wrembel's (2012) study suffered from problems with the status of L3 English (i.e. in terms of chronology it was the subjects' L2, in terms of dominance—L3, with more dominant French as L2), which could have affected the findings. The present results seem to be consistent with the assumption of a combined cross-linguistic influence that involves simultaneous influence of more than one previously acquired languages on the target language (cf. De Angelis 2007). This claim, posited from the perspective of Third Language Acquisition, constitutes an extension of a traditional SLA view of cross-linguistic influence (CLI), which has been perceived to be of a one-to-one type between the source and the target language. The present study provides some further evidence for the existence of a combined CLI, although not to the extent found in previous L3 studies.

The fourth hypothesis stipulated that there are differences in accent ratings due to the task and inter-rater variability. Although the inter-rater reliability proved very high in the four rating tasks (i.e. accent, intelligibility, certainty, acceptability), there was less consistency in the identification of the subjects' L1 and this

task turned out to be particularly difficult for the raters. Moderate to high correlations were also observed in all four parameters in case of the inter-speaker ratings, i.e. the scores for the read and spoken performance of the same subjects were correlated to a considerable extent.

Analysis of the inter-rater variability led to the conclusion that the native speaker status and the length of teaching experience exert some influence on foreign accent ratings. The findings did not follow the trends reported in some previous accent studies, in which native raters assigned higher points for accent and non-natives proved to be more severe judges (cf. Scheuer 2000; Wrembel 2010, Wrembel under review). To the contrary, they demonstrated the opposite tendencies, although the differences between native versus non-native judgements were significant only for intelligibility scores. The non-native raters found the subjects more intelligible most probably due to the shared L1 Polish. On the other hand, the native French raters were found to be significantly more confident in their ratings than the non-native judges. Furthermore, there was also an effect of the length of EFL teaching experience on the confidence rating. On the whole, although the findings were fairly consistent when it comes to the patterns of ratings, the results displayed some variability with respect to the raters' nativeness status and their teaching experience.

5 Conclusions

The contribution was aimed to shed a new light on the hotly debated nature of cross-linguistic interference in third language phonological acquisition. This foreign accent study on the L1 Polish/L2 English/L3 French language combination was expected to validate previous findings by the same author by providing a mirror design study to that of Wrembel (2012) (i.e. by exchanging the L2/L3 language combinations from L1 Polish, L2 French, L3 English) and by focusing on typologically less related languages (French and English) as opposed to the typologically related combination (i.e. L1 Polish, L2 German, L3 English) which was investigated in Wrembel (2010).

The present results, unlike those reported in Wrembel (2010) and Wrembel (2012), neither support nor fully disconfirm Hammarberg and Hammarberg's (1993, 2005) claim about the supremacy of the non-native influence at the initial stages of L3 phonological acquisition. The transfer from the subjects' L1 Polish phonology dominated in the subjects' performance in L3 French irrespective of their language proficiency in L3 French, however, some L2 English influence was also evidenced based the raters' identification of the subjects' L1.

The findings are consistent with the assumption of a combined cross-linguistic interference (cf. De Angelis 2007) thus substantiating the existence of both native and non-native influences in L3 phonology, although their relative strength and prevalence is hypothesised to be dependent on the typological relatedness of the respective language combinations. Trying to account for the discrepancy in the

results between the present study and that Wrembel (2010) or Wrembel (under review), one may point to an important role played by the factor of typological distance between the language repertoires under investigations (cf. the Typological Primacy Model by Rothman, 2011). The strength of the L2 status effect on the L3 interphonology was more evident in the accent rating study on typologically related L2 English and L3 German (Wrembel 2010) than in the present study involving typologically more distant L2 English and L3 French.

Future research aims to broaden the research perspective provided by L3 accent ratings by investigating selected acoustic phonetic dimensions such as VOT patterns or vowel duration as measures of accentedness in the acquisition of L3 phonology.

References

Cabrelli Amaro, Jennifer and Jason Rothman. 2010. On L3 acquisition and phonological permeability: A new test case for debates on the mental representation of non-native phonological systems. *IRAL* 48, 275–296.

Carlisle, Robert. 1991. The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied Linguistics* 12/1, 76–95.

Cenoz, Jasone and Ulrike Jessner (eds). 2000. *English in Europe: the Acquisition of a Third Language*. Clevedon: Multilingual Matters.

Cenoz, Jasone, Britta Hufeisen, Ulrike Jessner (eds.) 2001. *Cross-linguistic Influence in Third Language Acquisition*. Clevedon: Multilingual Matters.

Chamot, Anna 1973. Phonological problems in learning English as a third language. *International Review of Applied Linguistics* XI: 3, 243–50.

De Angelis, Gessica. 2007. *Third or Additional Language Acquisition*. Clevedon: Multilingual Matters.

Gallardo del Puerto, Francisco, Esther Gómez Lacabex and Maria Luisa García Lecumberri. 2007. The assessment of foreign accent by native and non-native judges. *PTLC Proceedings*, London, CD-ROM.

Gut, Ulrike. 2010. Cross-linguistic influence in L3 phonological acquisition. *International Journal of Multilingualism* 7 (1): 19–38.

Hammarberg, Björn. 2001. Roles of L1 and L2 in L3 production and acquisition. In *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives,* eds. Jasone Cenoz, Britta Hufeisen, Ulrike Jessner, 21–41, Clevedon: Multilingual Matters.

Hammarberg, Björn and Britta Hammarberg. 1993. Articulatory re-setting in the acquisition of new languages. *Phonum* 2: 61–67.

Hammarberg, Björn and Britta Hammarberg. 2005. Re-setting the basis of articulation in the acquisition of new languages: A third-language case study. In *Introductory Readings in L3*, eds. Britta Hufeisen and Robert Fouser, 11–18, Tübingen: StauFFenburg Verlag.

Højen, Anders. 2000. Detection of Danish accent in English by native speakers of Danish. In *Proceedings of 4th International Symposium on the Acquisition of Second-Language Speech, New Sounds 2000,* eds. Allan James and James Leather, 80–86, University of Klagenfurt.

Llama, Raquel, Walcir Cardoso and Laura Collins. 2010. The Influence of Language Distance and Language Status on the Acquisition of L3 Phonology. *International Journal of Multilingualism* 7 (1): 39–57.

Llisterri, Joaquim and Dolors Poch- Olivé. 1987. Phonetic Interference in Bilingual's Learning of a Third Language. In *Proceedings of the XIth International Congress of Phonetic Sciences*. Tallinn: Academy of Sciences of the Estonian SSR. 5: 134–147.

Piske, Thorsten, Ian R. A. MacKay and James Emil Flege. 2001. Factors affecting degree of foreign accent in an L2. A review. Journal of Phonetics 29. 191–215.

Pyun, Kwang-soo. 2005. A Model of Interlanguage Analysis—the Case of Swedish by Korean Speakers. In *Introductory Readings in L3,* eds. Britta Hufeisen and Robert Fouser, 55–70, Tübingen: Stauffenberg Verlag.

Ringbom, Håkan. 1987. *The Role of the Mother Tongue in Foreign Language Learning*. Clevedon: Multilingual Matters.

Rivers, Wilga 1979. Learning a sixth language: an adult learner's daily diary. *The Canadian Modern Language Review, 36*, 67–82.

Rothman, Jason. 2011 (to appear). L3 Syntactic Transfer Selectivity and Typological Determinacy: The Typological Primacy Model. *Second Language Research* 27 (1) Special L3 Issue edited by Rothman, Jason.

Scheuer, Sylwia. 2000. What makes foreign accent sound foreign? In *Proceedings of 4th International Symposium on the Acquisition of Second-Language Speech, New Sounds 2000,* eds. Allan James and James Leather, 306–314, University of Klagenfurt.

Wrembel, Magdalena. 2010. L2-accented speech in L3 production. *International Journal of Multilingualism.* Special Issue, 7 (1): 75–90.

Wrembel, Magdalena. (2012, in press). 'Foreign accentedness in Third Language Acquisition; the case of L3 English', in Jennifer Cabrelli Amaro, Susan Flynn and Jason Rothman (eds) *Third Language Acquisition in Adulthood*. Amsterdam: John Benjamins.

Wunder, Eva. 2010. Phonological cross-linguistic influence in third or additional language acquisition. In *Proceedings of 6th International Symposium on the Acquisition of Second-Language Speech, New Sounds 2010*, eds. Katarzyna Dziubalska-Kołaczyk, Magdalena Wrembel and Małgorzata Kul, 566–571, CD-ROM version, Poznań.

Acquiring Angma: The Velar Nasal in Advanced Learners' English

Wiktor Gonet, Jolanta Szpyra-Kozłowska and Radosław Święciński

Abstract The English velar nasal, known as *angma* or *engma*, belongs to consonants that are most difficult to master by Poles, in spite of the fact that this sound occurs also in Polish as a result of the assimilation of the dental nasal to the following velar plosives (e.g. in *tango* 'tango' pronounced as [ˈtaŋgɔ]). It is problematic for Polish learners, however, in other contexts, i.e. word-finally, as in *bring*, before vowels, as in *singer*, and before non-velars, as in *strongly*, where the velar nasal is pronounced with a following velar plosive. This paper examines the acquisition of the velar nasal by 60 advanced/proficient Polish learners of English, students at the English Department of Maria Curie-Skłodowska University, Lublin, Poland. The realization of angma is scrutinized with a view to uncovering regularities in its acquisition by the participants, establishing their success rate and the degree of difficulty in the production of the velar nasal in each of the three problematic contexts. The obtained results are compared with those pertaining to students' acquisition of ash, schwa and unstressed unreduced vowels in order to characterize advanced learners' interlanguage in more detail. The presented observations carry important pedagogical implications for the phonetic training of Poles.

1 Introduction

While much research has been concerned with the acquisition of English pronunciation by beginning and intermediate Polish learners, relatively little is known about advanced and proficient learners' interlanguage (see, however,

W. Gonet (✉) · J. Szpyra-Kozłowska · R. Święciński
Institute of English, Maria Curie-Skłodowska University, Lublin, Poland
e-mail: gonet@umcs.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_4, © Springer-Verlag Berlin Heidelberg 2013

Nowacka 2008). In a series of studies (Gonet et al. 2010a, b, c) we have undertaken an examination of phonetic progress and ultimate achievement of English Department students representing a high level of language proficiency in selected aspects of English phonetics particularly problematic for Poles. Thus, we have scrutinized their acquisition of unstressed unreduced vowels, ash ([æ]) and schwa [ə]), paying particular attention to the degree of difficulty involved in the realization of these segments in different phonological (segmental and prosodic) contexts.[1] This study is a continuation of our earlier research, this time with the focus on the velar nasal.

As noted by other researchers, e.g. Sobkowiak (1996: 94), "*the velar nasal is among the hardest to master by Poles*" in spite of the fact that such a consonant is found in Polish, both in native words, e.g. *ręka* 'hand' [ˈrɛŋka], as well as in borrowings, e.g. *tango* 'tango' [ˈtaŋgɔ]. The problem with this segment lies in the fact that, as shown in the above examples, in Polish it occurs exclusively in the context of the following velar plosives while in English it has a wider distribution in that it is also found word-finally, as in *bring,* before vowels, as in *singer,* and before non-velars, as in *strongly.* It is exactly in these three contexts that the production of the consonant in question is problematic for Polish learners of English. A particularly difficult case concerns the items with two occurrences of [ŋ] in one word, as in *singing*.

As a matter of fact, the issue concerns not so much the articulation of angma itself, as the suppression of the following plosive added to it due to two combined factors: distributional restrictions on the occurrence of this segment in Polish, reinforced by a powerful impact of the English spelling rules. Thus, typical of English spoken by Poles are forms such as [ˈlɔŋk], [ˈsiŋgə], [ˈstrɔŋgli], [ˈriŋgiŋk][2] with the resulting impression of a strong foreign accent (Gonet 1982; Avery and Ehrlich 1992).

In this paper we examine the acquisition of the velar nasal in the problematic contexts by advanced to proficient students of English. More specifically, the following issues are addressed:

- the participants' progress and success rate in the acquisition of angma;
- the degree of difficulty in mastering the velar nasal in each of the problematic contexts;
- pedagogical implications for the phonetic training of advanced Polish learners of English.

[1] In this respect our studies differ from Nowacka's (2008) work in which Polish students' phonetic progress is examined without, however, taking into account contextual factors.

[2] The presence of [k] in some items is due to Polish word-final obstruent devoicing as well as to regressive voice assimilation in obstruent clusters.

2 Experimental Design

Sixty students, both male and female, of the Maria Curie Skłodowska University English Department, took part in the experiment. They formed three groups which consisted of twenty Year 1, Year 3 and Year 5 students, representing, roughly, upper-intermediate, advanced and proficient learners, respectively. Within each year, the participants were selected randomly. The sample tested satisfactorily for homogeneity.

Sixty English lexical items were selected constituting four groups of 15 words with angma occurring in different contexts: in word-final position (e.g. *among*), before non-velars (e.g. *kingdom*), before vowels (e.g. *singer*) and with two angmas, one medial before a vowel, and one final (e.g. *singing*). A list of all the test items is placed in Appendix 1. The test words were used in 15 diagnostic sentences (see Appendix 2).

The participants, unaware of the experimental goals, were given a list of the diagnostic sentences to study for several minutes, and then were asked to read them aloud at their own pace. Their performance was individually recorded, yielding 3600 tokens. The data thus obtained were analysed auditorily and assessed by a panel of experts consisting of 3 university pronunciation teachers who evaluated the recordings independently. In cases of disagreement, the tokens in question were replayed and discussed. Where the agreement still could not be reached, the experts resorted to slow-rate replay and, if necessary, to spectrographic visualizations. All doubts were finally resolved in this manner.

3 Results and Discussion

The focus of this section is on the students' phonetic progress and their final attainment in the production of angma as well as on examining the realizations of this consonant in different phonological contexts. Hence the two independent variables studied here are Years and Contexts, while the dependent variable is the correctness score, or index, expressed as percentage of correctly pronounced tokens out of all potentially possible correct pronunciations.

3.1 Progress and Ultimate Achievement in the Production of Angma

38 % of all the tokens containing angma were pronounced properly, while 62 % were realized incorrectly.[3] Figure 1 shows the composition of the correctly pronounced items when split into Years.

[3] Incorrect realizations included the following: [ŋg, ŋk, ndž, n, ng, nk].

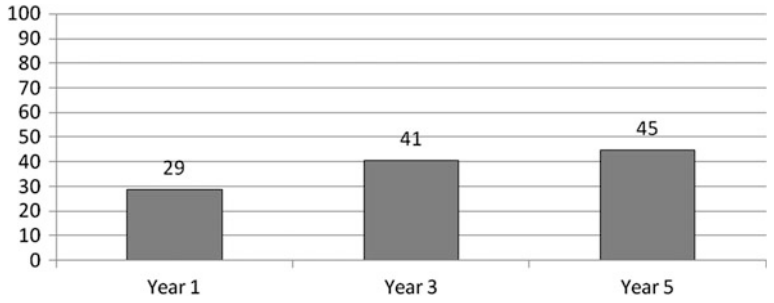


Fig. 1 The split of the correct score into years: Main Effect of Years

Figure 1 demonstrates that Year 1 students' correctness score was 29 %, that of Year 3–41 %, and that of Year 5–45 %. The 12 % progress between Years 1 and 3 was statistically significant at $\alpha = 0.001$, while a further 4 % advancement between Years 3 and 5 was not significant. Naturally, the progress from Year 1 to Year 5 of 16 %, was significant at the same level. The final outcome of 45 % correct realizations of the velar nasal by the 5th year students was assessed as highly unsatisfactory regarding the pronunciation proficiency expected of the subjects.

3.2 Angma in Different Phonological Contexts

The second goal of this study was to examine whether there exists a meaningful connection between the degree of difficulty in producing angma and four studied contexts in which the sound was difficult to pronounce for Polish learners: word-finally, before non-velar consonants, before vowels and the double occurrence within a lexical item. A plot of Main Effect of Context summarizes the relevant data (Fig. 2).

Thus, viewed from left to right, angma was most problematic between vowels, with an average correctness score equaling 16 %, less so in words with two angmas (32.3 %), still easier if it occurred before non-velars (41.3 %), and relatively easy to pronounce in word-final position (62.7 %). All the differences between the means presented in Fig. 2 were significant at = 0.001, showing that the studied contexts exerted a significant effect upon the precision of articulation. Let us now split each of the Main Effect means into its component context means; the result is shown in Fig. 3.

Figure 3 shows that the correctness scores for Years patterned differently across the four examined contexts. Thus, going from left to right across Fig. 3, the intervocalic angmas were most difficult to pronounce, with no significant differences between the years. In the case of double angmas, the results scored by Year 5 were the highest (39 %) and, quite surprisingly, differed statistically from the

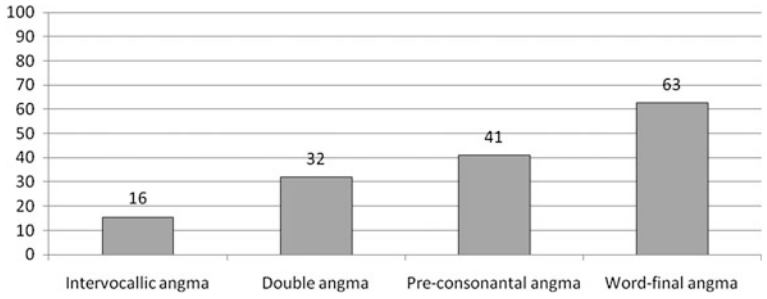


Fig. 2 Main Effect of Context showing the correctness indices of angma averaged across Years (in %)

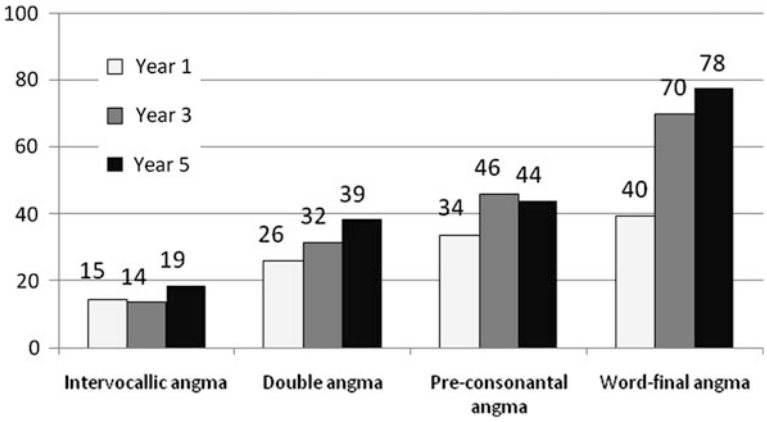


Fig. 3 Correct realizations of angma in 4 contexts (in %)

other two Years at $\alpha = 0.001$, although the 1 % smaller difference between Year 3 and Year was evaluated as insignificant even at $\alpha = 0.05$. In pre-consonantal angmas, a statistically significant difference was inferred between Year 1 and Year 3, also at $\alpha = 0.001$, while the difference between Year 3 and Year 5 was insignificant at $\alpha = 0.05$. In fact, the performance of Year 5 slightly lowered when compared to Year 3. Word-final angmas scored the highest correctness indices, showing a statistically significant progress from Year 1 to Year 3 (at $\alpha = 0.001$), with no significant difference between Year 3 and Year 5, even when tested at $\alpha = 0.05$.

These data can also be interpreted dynamically as the rate of progress between the successive levels of Years. Thus, the students' performance with respect to intervocalic angmas did not exhibit any marked progress across Years (the means increased only by 4 %), and in words with double angmas the overall progress was also small (13 %). A 10 % improvement was attested in pre-consonantal (i.e. before non-velars) angmas, while the rise in word-final angmas was the strongest of all, and amounted to 38 %. It seems that this much higher success rate in the production of

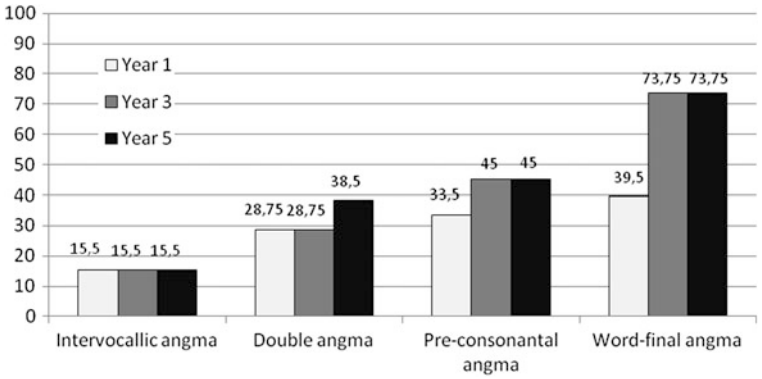


Fig. 4 Contextual realizations of angma with averaged non-significant differences (in %)

word-final nasals stems both from a very high frequency of -ing forms in English, as well as from the focus of phonetic instruction on this particular context. Another interesting observation is that angma is easier to pronounce before vowels when another such consonant is found at the end of the same word, as in *singing*.

The significance relationships described above can be more readily comprehended if shown on a new illustration in which the non-significant differences were averaged (cf. Fig. 4).[4]

Looking at the data above, one can assume that the phonetic training that the students received during the 1st and 2nd year of studies was insufficient to boost their proficiency in pronouncing the velar nasal intervocalically and in words with two angmas (as mentioned above, one of the angmas in such words appeared intervocallically). The pre-consonantal and word-final angmas, however, were realized with marked improvement by Year 3 and Year 5, which may be attributed to the practical phonetics course that the participants attended during the first two years of studies.

As indicated by the Main Effect means, Context also exerted an effect on the rate of progress in the correctness of the pronunciation of the examined sounds; this is best shown in Fig. 5, where the Main Effect of Years (cf. Fig. 1) is split into its component contexts means.

Let us discuss the data moving, again, from left to right. Within each Year, the same relations among contexts are preserved: for every Year, the most difficult to pronounce is the intervocalic angma; words containing double angmas are easier to pronounce, and pre-consonantal angmas are still more so. The easiest to pronounce are word-final angmas, especially in Year 3 and Year 5.

[4] This procedure was first described in Gonet 1989.

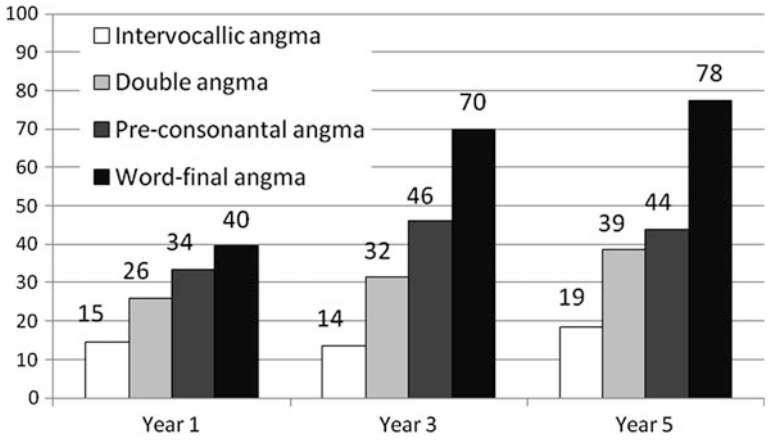


Fig. 5 Contextual realizations of angma within each Year (in %)

3.3 Problems with Angma Versus Difficulties in the Pronunciation of Vowels

Comparing the data with the results obtained in previous experiments on the students' acquisition of unstressed unreduced vowels, ash and schwa (see Gonet et al. 2010a, b, c), one can observe a similar phonetic proficiency level with about 30 % of correct realizations in Year 1 students (cf. Fig. 2). In the case of ash and schwa, however, a marked overall progress of 33 % (Gonet et al. 2010a) and 36 % (Gonet et al. 2010b) could be observed between Year 1 and Year 5 (with a similar final achievement of about 70 % in Year 5), whereas in the case of angma the improvement amounts to 16 % only.

Thus, contrary to a common belief that mastering English consonants is relatively easy in comparison with vowels, the results we obtained show that advanced Polish learners of English find proper production of the velar nasal in the four examined contexts more problematic than the articulation of the two vowels (Fig. 6).

This interesting regularity can surely be attributed to the presence of intensive phonetic training in the curriculum within the first two years of the study.

Another observation made in our previous studies (Gonet et al. 2010a, b, c) that finds further support in the present work is that the dependence on the phonological context of the students' realization of problematic sounds is strongly related to their level of proficiency. Thus, for the least advanced freshmen, all contexts appear to represent a similar degree of difficulty, but starting with Year 3 and continuing with Year 5, the sound's segmental and/or prosodic environment begins to play an increasingly decisive role. Moreover, in none of the instances that have been scrutinized is the dependence between the phonological context and the level correctness stronger than it is in the case of angma.

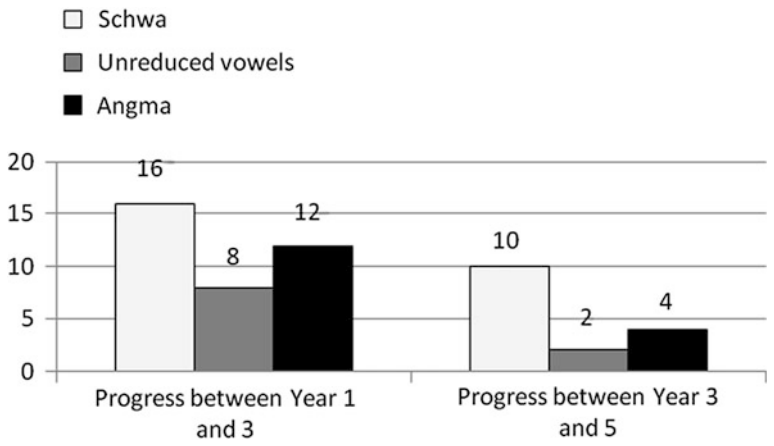


Fig. 6 Comparison of progress rate between Years 1, 3 and 5

4 Conclusions

This study, while focusing on the intricacies involved in the production of the velar nasal, also sheds light on the development of other pronunciation skills in Polish university students of English. It also characterizes certain important properties of these learners' interlanguage.

In all the examined instances a steady phonetic progress in the realization of problematic sounds was observed within the five years of students' university education. However, most of this progress took place within first two years of the study. In the case of angma, statistically significant progress was attested only between Years 1 and 3.

Our experiment also uncovered a strong effect exerted by the phonological context in which the velar nasal occurs, with the following order of increasing difficulty: word-finally, in words with two velar nasals, before non-velars and before vowels. The presence of two angmas in a single word resulted in an improved production of intervocalic velar nasals, which can be attributed, somewhat paradoxically, to the especially high degree of difficulty of the cases that require an extra articulatory effort on the learner's part.

Another interesting observation can be made here concerning context-dependence in the realization of English sounds and learner proficiency: the more advanced they are, the stronger is the effect of the phonological environment in which a sound is found.

Pedagogical implications of this research are obvious: proper phonetic training of advanced and proficient learners should continue throughout the whole period of their studies, focusing especially on the difficult phonological contexts in which the practiced sounds occur. In the case of the velar nasal, a more thorough training appears necessary, particularly when it occurs between vowels.

Appendix 1

Appendix 1 List of lexical items used in the experiment

	Word-final	Before non-velars	Before vowels	Sequence of 2 angmas (before a vowel and final)
1.	Among	Kingdom	Springer	Singing
2.	Nothing	Hangman	Hangover	Banging
3.	Along	Belonged	Youngish	Hanging
4.	Drinking	Songs	Longish	Springing
5.	Looking	Springfield	Singer	Ringing
6.	Wrong	Pangs	Hangout	Flinging
7.	Rang	Wellington	Birmingham	Upbringing
8.	Reading	Strongly	Buckingham	Prolonging
9.	Ring	Things	Bingen	Clinging
10.	Gang	Langdale	Coat hangers	Belonging
11.	Playing	Arlington	Ringers	Swinging
12.	Ping	Strings	Nottingham	Slinging
13.	Pong	Earrings	Slangy	Stinging
14.	Horse-riding	Surprisingly	Stringy	Stringing along
15.	Young	Wrongly	Banger	Wringing

Appendix 2

List of Sentences used in the experiment

1. In the old days in most *kingdoms* the *hangman's* job *belonged* to highly prestigious professions.

2. She kept *wringing* her hands in despair when she thought of the way that *gang* was *stringing* her husband *along*.

3. When he lived *among* the bushmen there was *nothing* he missed more than "Jerry *Springer* Show."

4. After the night of *drinking*, *singing songs* and *banging* on the drums in a *Nottingham* bar they had a terrible *hangover* and *pangs* of conscience.

5. This *youngish-looking* man with *longish* hair was a *singer* in a terrible *hangout* in *Wellington* suburbs.

6. With his aristocratic *upbringing* he felt quite at ease in *Buckingham* Palace, where he received a medal and a royal *ring*.

7. I *strongly* agree that it was *wrong* to allow the children to do all that *hanging* from the ropes, *springing* from the furniture, *ringing* the bells and *flinging things* at each other.

8. They *rang* us up to tell us about their weekend trip to *Reading, Langdale, Birmingham* and *Arlington*.

9. *Surprisingly* we didn't know they enjoyed *horse-riding* and *playing ping pong*.

10. He thought of *prolonging* his stay in *Bingen* in order to buy new *strings* for his guitar, nice *earrings* for his girl-friend and some wooden *coat hangers*.

11. Because of the fire in our neighbours' house our eyes were *stinging* from the smoke.

12. The *ringers* of the bells kept *clinging* to the ladder of the church tower.

13. This *young* man from *Springfield* had *stringy* arms and *slangy* speech.

14. He drove his old *banger swinging* it from the left to the right and *slinging* mud all over the place.

15. They were *wrongly* suspected of *belonging* to a terrorist organization.

References

Avery, Peter, Susan Ehrlich. 1992. *Teaching American English Pronunciation.* Oxford: Oxford University Press.

Gonet, Wiktor. 1982. 'Legato speech', or the psychology of teaching phonetic fluency. *Lubelskie Materiały Neofilologiczne*, pp. 307–319.

Gonet, Wiktor. 1989. *Factorial Analysis of the Duration of R. P. Monophthongs in Monosyllabic Words.* Unpublished Ph. D. dissertation. Lublin: UMCS.

Gonet, Wiktor, Jolanta Szpyra-Kozłowska, Radosław Święciński. 2010a. Clashes with ashes. In Waniek-Klimczak, Ewa (ed.), pp. 213–232.

Gonet, Wiktor, Jolanta Szpyra-Kozłowska, Radosław Święciński. 2010b. The acquisition of Vowel Reduction by Polish students of English. In Waniek-Klimczak (ed.), pp. 291–308.

Gonet, Wiktor, Jolanta Szpyra-Kozłowska, Radosław Święciński. 2010c. Unstressed unreduced English vowels in the speech of Polish students—a developmental study. In Waniek-Klimczak (ed.), pp. 325–340.

Nowacka, Marta. 2008. *Phonetic Attainment in Polish University and College Students of English. A Study in the Productive and Receptive Pronunciation Skills.* Unpublished Ph.D. dissertation. Maria Curie-Skłodowska Uniersity, Lublin, Poland.

Sobkowiak, Włodzimierz. 1996. *English Phonetics for Poles.* Poznań: Bene Nati.

Waniek-Klimczak, Ewa. (ed.). 2009. *Issues in Accents II: Variability and Norm.* Cambridge Scholars Publishing.

Vowel Quality and Duration as a Cue to Word Stress for Non-native Listeners: Polish Listeners' Perception of Stress in English

Arkadiusz Rojczyk

Abstract The paper reports results from a study on the perception of vowel quality and duration as a stress cue in English by Polish learners of English. The word *record* was synthesised in which F0 was held constant and vowel quality and duration were manipulated to obtain three different types of interaction. The two parameters could remain neutral as to the stress location, contribute jointly to stress location, or provide conflicting cues. Thirty-two Polish advanced learners of English participated in an identification task in which they were required to choose a noun form *'record* stressed on the first syllable or a verb form *re'cord* stressed on the second syllable. The results revealed that in the absence of f0 cues, the listeners were able to make use of both vowel quality and duration to judge the stress location. It is hypothesized that, in the absence of the cue which is primary in their native language, the listeners were able to resort to cues which are secondary in their native language when listening to L2 speech.

1 Introduction

Effective communication in a non-native language requires the processing of phonetic cues which are typical for the implementation of word stress in that language. Native English listeners use word stress to segment speech and locate word boundaries (Cutler 2005; Cutler and Norris 1988; Tyler and Cutler 2009). Non-native listeners may experience a degradation in performance due to a different weighting of word-stress cues in L2 compared to their native language

A. Rojczyk (✉)
University of Silesia, Katowice, Poland
e-mail: arkadiusz.rojczyk@us.edu.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_5, © Springer-Verlag Berlin Heidelberg 2013

(Cooper et al. 2002; Cutler 2009; Cutler et al. 2007). Cues that are robust for native listeners may be largely unattended by non-native listeners and cues less relevant in native perception can be given too much weight in non-native perception.

Non-native perception of stress may be investigated from two different perspectives: phonological and fine-grained phonetic. Phonologically Polish uses a metrical structure that predominantly assigns stress to a word's penultimate syllable (e.g., Dłuska 1947; Comrie 1967; Wierzchowska 1980).[1] Using a metric of "stress deafness" (Dupoux et al. 2001, 2008; Peperkamp and Dupoux 2002), which predicts the difficulty in cross-linguistic perception of stress based on the listeners' phonological representations in L1, Peperkamp et al. (2010) classified speakers of Polish as exhibiting an intermediate patterns of "stress deafness". In a sequence recall task, Polish listeners performed better than French listeners, whose native language has fixed stress, and worse than Spanish listeners, whose native language has an irregular stress pattern. This result is taken as evidence that lexical tendencies typical of a given language influence cue sensitivity. The more predictable the stress position in a word, the less encoded it is and the less cue sensitivity is necessary to process stress. The difficulty for Polish speakers in correct in identification of stress in languages with much less predictable metrics is also demonstrated by studies investigating how Polish learners of English compute a stress pattern in English (Archibald 1992; Waniek-Klimczak 2002). Polish learners use complex strategies to assign stress in English words which is a product of both L1 transfer and the effects of Universal Grammar. They tend to compute stress by generalising syntactic categories of words. They are also quantity sensitive, in that they assign stress to syllables containing long vowels or diphthongs.

More relevant to the current study is the production of fine-grained phonetic parameters that contribute to the perception stressed and unstressed syllables. In the following sections, we discuss the contribution and perceptual hierarchy of such parameters in English and Polish. We also look at how a combination of stress cues typical for English is acquired and processed by non-native speakers.

2 Acoustic Correlates of Stress in English

Lexical stress in English is acoustically manifested by combining cues such as frequency (f0), intensity, vowel duration, and vowel quality (e.g., Beckman 1986; Fry 1955, 1958; Lieberman 1960; Sluijter and van Heuven 1996). Stressed syllables are produced with some combination of higher F0, greater intensity, longer vowel durations, and more peripheral vowel quality, though not all features need be present. Inappropriate realisation of any of those parameters leads to poorer perception of

[1] This stress patterns is largely predominant but not unexceptionally invariant. Peperkamp et al. (2010) analysed two corpora of spoken Polish, "Szuman" (Smoczyńska 1985) and "Weist" (Weist et al. 1984), and found that around 0.1 % of the words have an exceptional stress pattern.

stress by native listeners (Zhang and Francis 2010; Zhang et al. 2008). For example, Beckman and Edwards (1994) analysed recordings of a sequence /ˈpapa/ in three different speech rates, controlling the utterance position of the target. They found that stressed syllables had longer durations and their opening movement was bigger and faster. Crystal and House (1988) reported English monophthongs to be 50 % shorter in unstressed compared to stressed syllables. Fear et al. (1995) recorded British English speakers' pronunciations of words such as *audiences*, *auditoria*, *addition*, and *audition*, in which the first syllable had different degrees of stress and vowel reduction. Acoustic analyses revealed that duration, intensity, and spectral tilt were significantly different in stressed and unstressed syllables. Others studies have demonstrated f0 movement and vowel quality differences in stressed and unstressed syllables in words such as *absence* vs. *absurd* (Braun et al. 2008) and in a corpus of spoken English (Yuan et al. 2008).

Establishing a relative weighting of spectral cues in perception of stress in English is difficult because stress is used to convey prominence not only at the level of word stress but also on higher levels, such as phrase stress and sentence stress (Beckman and Edwards 1994; Braun et al. 2011; Ortega-Llebaria et al. 2010). In one of the earliest studies, Fry (1958) synthesized noun–verb word pairs that were presented to the English-speaking listeners. The analysis of judgments pointed to a hierarchy of cues which placed F0 movements as the most robust stress cue, followed by vowel duration. The effect of intensity was negligible. The weak effect of intensity relative to other parameters was also observed in other studies (Mol and Uhlenbeck 1956; Sluijter and van Heuven 1996; Turk and Sawusch 1996).

Most of the research on the acquisition of English word stress by non-native learners has concentrated on vowel quality and duration. The subset of unstressed vowels which are not realised as a schwa in English are characterized by qualitative reductions, in that they attain midrange frequencies of the first two formants compared to their stressed counterparts. This systematic variability of quality has been demonstrated to play a great role in cueing stressed syllables in English (Beckman 1986; Campbell and Beckman 1997; Cooper et al. 2002; Cutler 1986; Gay 1978; Lehiste 1970). Huss (1978) found that vowel quality, together with F0 contour, outranked both duration and intensity in a noun–verb perception experiment. Fear et al. (1995) used cross-spliced words from the production study reported earlier in a naturalness judgment task. Their English listeners were most sensitive to the words in which stress position did not agree with vowel quality, which was taken as evidence that English listeners are primarily sensitive to a change in vowel quality as a stress cue. Cutler and Clifton (1984) reported that hearing an unreduced full vowel in an unstressed position or a reduced vowel in a stressed position produces longer RTs in a speeded semantic decision task. More recent studies confirm the robustness of vowel quality in stress judgments by speakers of English. Tyler and Cutler (2009) used artificial-learning experiments to show that English listeners locate word boundaries at the onset of a strong syllable, that is the one containing a full vowel. Zhang and Francis (2010) manipulated a naturally recorded noun–verb pair *desert* to obtain stimuli differing in F0,

intensity, duration and vowel quality. They reported that in most cases vowel quality was more salient than the other cues. Yu and Andruski (2010) tested both identification and discrimination of stressed syllables consisting of real words, pseudowords, and hums. They observed that English speakers make use of vowel quality cues especially in pseudowords, when lexical information is not available. Finally, Braun et al. (2011) reported that inappropriate reduction of vowel quality in unstressed syllables impedes English listeners' perception of Dutch-accented English.

Durational variability in stressed and unstressed syllables is manifested in the tendency of vowels to be longer in stressed and shorter in unstressed positions. However, the relative contribution of duration to the perception of stress in largely unresolved. Fry (1958) showed that, although duration alone can trigger perception of stress, it is fully effective only in trading relations with intensity. Huss (1978) reported that duration is an insufficient cue to words stress in English. Campbell and Beckman (1997) did not find duration to be a consistent cue in English. Other studies confirmed the contribution of duration to stress perception. Results reported by Zhang and Francis (2010) pointed to duration as a cue in a lexical stress judgment task, yet vowel quality was more salient. Yu and Andruski (2010) also found duration as a consistent stress cue for English listeners. Tyler and Cutler (2009) reported that durational variability related to stress is actively used in speech segmentation. Yet another study (Beckman 1986) ranked duration high in the hierarchy of perceptual robustness, even higher than vowel quality.

3 Acoustic Correlates of Stress in Polish

Polish is classified as a stress-timed language with no phonological vowel reduction (Crosswhite 2003; Jassem 1962). Jassem (1962) failed to observe any vowel reduction in Polish and concluded that F0 is the only correlate of stress. The stability of Polish vowels under different stress conditions is confirmed in a typology of perception of rhythm classes by Ramus et al. (2003). Dogil (1995, 1999) tested Jassem's (1962) predictions using a constructed dialogue paradigm and arrived at similar conclusions pointing to the frequency and trajectory of F0 as the only reliable correlate of word stress in Polish. Yet another study reported variability of spectral tilt in stressed and unstressed syllables in Polish (Crosswhite 2003), however, it is not made clear how perceptually robust this parameter is relative to F0.

Although the fact that Polish does not reduce vowels phonologically in unstressed syllables is generally accepted, some recent research has concentrated on measuring vowels in unstressed positions using more precise research designs. All those studies were driven by predictions from the undershoot model (Lindblom 1963; see also Gay 1978) that decreased durations of vowels in unstressed positions would fail to provide a sufficient time-span for articulators to reach their target positions. The transitions from one to the other flanking consonant will be so

rapid that a midpoint of a vowel is likely to be undershot relative to its target. Strycharczuk and Jurgec (2008) analysed selected words from the online dictionary of Polish and observed lower F2 values for front vowels in unstressed positions compared to their counterparts in stressed positions and monosyllabic words. Nowak (2006) reported that Polish unstressed vowels exhibit reduction in the F1 domain as well as in the F2 domain, however, for the latter the evidence was less clear. Jaworski (2009) provided measurements from one speaker of Polish that showed qualitative reductions of unstressed vowels in fast speech. Finally, Rojczyk (2010a) recorded 23 Polish speakers' productions of /dVdVdV/ sequences in a reiterant speech mode. The second syllable was invariably stressed in accord with penultimate-stress metrics in Polish. Both the unstressed (initial) and stressed vowels were measured. The results revealed significant lowering of F1 values in unstressed positions for all vowels except for /ɨ/ and significant rising of F2 values for back vowels.

Vowel duration differences in Polish stressed and unstressed syllables were reported in an early study by Jassem (1962), who estimated that stressed vowels are 17 % longer than unstressed vowels. Nowak (2006) found an 18 % duration difference across various speech styles. Rojczyk (2010a) reported a 30 % difference in duration between stressed and unstressed vowel. This ratio is observably larger relative to the two previous studies, but may be explained by specificity of a reiterant speech mode, which was used to elicit speakers' productions.

Although studies noted earlier point to some vowel reduction in Polish, its magnitude is much less pronounced than in English. Polish reductions are not coded phonologically and seem merely to be a by-product of speech kinesthetics in general. This assumption is further supported by research showing difficulties that Polish learners of English have when acquiring vowel reduction in English.

4 English Vowel Quality and Duration as a Stress Cue in Non-Native Speech

Non-native speakers' inability to reduce vowels in English unstressed syllables has been argued to contribute importantly to detectable foreign accent (Flege and Bohn 1989; Fokes and Bond 1989), especially in lexical stress production tasks (Fokes et al. 1984). This pattern is transferred to perception as demonstrated by the results (Cutler et al. 2007; Sluijter et al. 1997; Tyler and Cutler 2009) showing that Dutch listeners will attend to suprasegmental cues rather than vowel reduction in speech segmentation.

The particular problems with correct productions and perception of English stress are influenced by typological characteristics of learners' native language. Most research has concentrated on languages that are so-called syllable-timed or tonic and thus have a different rhythmic organization from English. Braun et al. (2008) tested how Dutch speakers of English implement their native Dutch way of

producing lexical stress in English. They recorded pairs such as *absence-absurd* and found that their Dutch participants did not reduce qualitatively unstressed vowels in English. In a later study (Braun et al. 2011), they used the same recordings in a cross-modal priming task to investigate the perception of Dutch-accented speech by English listeners. They found that recognition suffered considerably due to insufficient reductions of vowels in unstressed syllables. Spanish speakers of English have been characterised by exaggerated durations of vowels in stressed syllables and insufficient qualitative reductions in unstressed positions (Flege and Bohn 1989; Hammond 1986). Similar observations were made for Japanese speakers of English (Anderson 1993; Kondo 2000; Mochizuki-Sudo and Kiritani 1991, but see Lee et al. 2006). Zhang et al. (2008) measured Mandarin speakers' productions of English two-syllable words differing only in stress position and found that their speakers diverged from native English controls in that they did not produce sufficiently reduced unstressed vowels. In a perception study (Zhang and Francis 2010), based on resynthesised tokens of *desert* (noun) and *desert* (verb), the listeners performed in a native like fashion in processing vowel duration and quality as a stress cue, which stands in contrast to the results obtained in the production study. In yet another study (Yu and Andruski 2010), Chinese listeners were found to use pitch as a primary cue for stressed syllables in English, while native English listeners used duration as the most consistent cue.

Research targeted specifically at Polish learners of English found predictable difficulties with qualitative and durational vowel reductions in English unstressed syllables. Gonet et al. (2010) found significant numbers of qualitatively unreduced vowels in Polish advanced learners' speech. This patterns was correlated with speakers proficiency, the most notable divergences from the native norm were observed for lower-level participants. Porzuczek (2008, 2010) reported insufficient durational reductions of English vowels in unstressed positions and function words measured in recordings of a test passage for Polish learners. Bogacka et al. (2006) found that they produced and perceived English unstressed /ə/ as a version of Polish /e/, which is much more peripheral in quality.

5 The Current Study

In the current study we aim to investigate how vowel quality and duration contribute to the perception of lexical stress in English by Polish advanced learners. The literature reviewed earlier points to a clear typological distinction between the fine-grained phonetic implementation of word stress in English and Polish. While English listeners make effective use of vowel quality and duration as markers of stress as well as using F0 cues, Polish listeners are predicted to rely mostly on F0 shifts. This prediction leads us to ask the following questions: how do Polish learners of English perceive word stress in English if the parameter of F0 is controlled and held constant? How does the variability in vowel duration contribute to the stress judgment? How does the quality of a vowel contribute to

the stress judgment? And finally, what is the interaction of vowel duration and quality if they are manipulated to interact with each other?

5.1 Participants

A total of 32 subjects participated in the study: 23 females and nine males. They were recruited from 2nd-year students at the Institute of English, University of Silesia. This selection guaranteed a uniform level of proficiency in English due to a regular administration of various tests in the use of English. They all volunteered and were not paid for their participation. None of the subjects reported any speech or hearing disorders.

The listeners participating in the study were selected from a larger pool by auditory and acoustic inspection of recordings from a pre-test recording session, in which they were asked to pronounce a difference between *record* (noun) and *record* (verb). The analysis revealed that eight participants did not produce the contrast for this pair and, as a consequence, they were eliminated from the listening experiment.

5.2 Stimuli

The MBROLA diphone synthesizer (Dutoit et al. 1996) was used to create a synthetic word *record*. The voice was male British English (en1 in a diphone database) set at 16100 with pitch 0.85. Each segment had a predefined duration of 116 ms and F0 was held constant at 97 Hz.

Next, the raw sequence was used as a baseline for manipulating duration and quality of the initial vowel in order to obtain stimuli differing in stress position. The initial vowel was synthesised into three different qualities: /e/ (F1 549 Hz; F2 1485 Hz; F3 2149 Hz). /ɪ/ (F1 369 Hz; F2 1831 Hz; F3 2357 Hz), and /ə/ (F1 445 Hz; F2 1321 Hz; F3 1909 Hz). Duration of the first and second vowel was manipulated by increasing the predefined duration of 116 ms by 50 % to 174 ms. Both vowel quality and duration were paired in such a way that they could either contribute jointly to signaling a stressed syllable, be used as a single parameter while the other was held neutral, or provide conflicting cues as to which syllable was stressed. This design was hypothesized to provide a general estimate of the relative weight of each parameter in the perception of English word stress by Polish learners. The stimuli may be tabulated as follows:

1. Initial /ɪ/ /rɪkɔ:d/
   (a) joint cues: /ɪ/ 116 ms; /ɔ:/ 174 ms
   (b) single cue: /ɪ/ 116 ms; /ɔ:/ 116 ms
   (c) conflicting cues: /ɪ/ 174 ms; /ɔ:/ 116 ms

2. Initial /ə/ /rəkɔ:d/

   (a) joint cues: /ə/ 116 ms; /ɔ:/ 174 ms
   (b) single cue: /ə/ 116 ms; /ɔ:/ 116 ms
   (c) conflicting cues: /ə/ 174 ms; /ɔ:/ 116 ms

3. Initial /e/ /rekɔ:d/

   (a) joint cues: /e/ 174 ms; /ɔ:/ 116 ms
   (b) single cue: /e/ 116 ms; /ɔ:/ 116 ms
   (c) conflicting cues: /e/ 116 ms; /ɔ:/ 174 ms

5.3 Procedure

The procedure was a self-paced identification task run from a Praat (Boersma 2001) script. There were nine stimuli × four repetitions randomized nonrepetitively for each listener. The experiment took place in the Acoustics Laboratory at the Institute of English, University of Silesia. Each session lasted approximately 15 min. The listeners were presented with orthographically rendered words 'a record' on the left and 'to record' on the right enclosed in rectangular shapes on a laptop screen. The 'play again' option was located in the middle of the screen and the participants were encouraged to use it whenever they deemed necessary. They were instructed to use a mouse to select the word that was the best representation of the stimulus they had heard. The stimuli were played binaurally over high-quality headphones at a comfortable listening level at 70 dB. The interval between a selection of a word on the screen and the presentation of the next stimulus was 1 s.

5.4 Analysis and Results

A factorial ANOVA design was used to analyse the results with vowel quality as a categorical factor (three levels: vowels /e/; /ə/; /ɪ/) and vowel duration as dependent variables (three measurements: equal duration in the two syllables; initial syllable longer; final syllable longer).

The effect of vowel duration irrespective of vowel quality was found to significantly influence the listeners' perception, $F(2, 190) = 3.8751$, $p < 0.05$. Stimuli with final longer vowel were identified more often as a verb *record*. The post hoc Fisher LSD showed that stimuli with final longer vowels were identified as stressed on the second syllable significantly more often than stimuli with normal-duration final vowels ($p < 0.01$). The lengthening of an initial vowels did not strengthen the perception of initial stress relative to initial vowels with normal duration ($p = 0.379$).

The analysis of vowel quality irrespective of vowel duration revealed a strong effect of an initial vowel type on stress judgments, $F(2, 93) = 30.267$, $p < 0.01$. Both initial /ɪ/ ($p < 0.001$) and /ə/ ($p < 0{,}001$) yielded more second-syllable stress judgments relative to initial /e/. The vowel /ɪ/ was more effective in cueing second-syllable stress compared to /ə/, however the effect was only close to significant ($p = 0.06$).

The detailed analysis of the interaction between vowel quality and duration showed that increased duration of a final syllable was dependent on vowel quality in the initial syllable. Lengthening of a final syllable strengthened the perception of a stress in a final syllable in the case of /ɪ/ ($p < 0.01$) but not for /e/ ($p = 0.09$) or /ə/ ($p = 0.88$).

6 Discussion

The results revealed that Polish advanced learners of English participating in this study were sensitive to both quality and duration in their stress judgments. Stimuli with longer final vowels were perceived as stressed on the second syllable. However, this effect was strongest for /ɪ/ relative to /ə/ and /e/. Vowel qualities /ə/ and /ɪ/ in the first syllable elicited a response of final stress. However, again, the effect was stronger for the latter relative to the former vowel.

The fact that vowel duration was used as a stress cue is interesting considering the fact that Polish does not use phonemic vowel duration contrasts. Moreover, Polish is a word-final devoicing language which neutralizes the voicing contrast in absolute codas. Although measurements have suggested some variability of vowel duration as a remnant of incomplete neutralization (Slowiaczek and Dinnsen 1985), perception experiments showed that this variability is not contrastively perceived (Slowiaczek and Szymanska 1989). This insensitivity of Polish listeners to temporal variability of a vowel was further supported in a study which demonstrated that Polish learners of English do not consistently produce (Waniek-Klimczak 2005) or perceive (Rojczyk 2010b) vowel duration as a cue to the voicing status of a following consonant in English. Other studies, however, provided evidence that vowel duration is used effectively by Poles as a subsidiary parameter in perception of English vowels (e.g., Rojczyk 2010c) Accordingly, our results may be interpreted as suggesting that vowel duration in English may or may not be a variable to Polish listeners depending on a task. It was found to be used here in the stress-judgment task.

Vowel quality was also used as a stress cue in the current study. The strongest effect was observed for vowel /ɪ/. We propose two different hypotheses as to why this might have been the case. The first hypothesis predictably relies on the tendency in English to use /ɪ/ in an initial syllable in a verb form *record*. This vowel, and not /ə/, is most often heard in native speakers' pronunciation. The listeners in our study may have learnt by statistical probability from the input to assign this vowel to an unstressed syllable in the verb *record*. The second hypothesis is based

on the assimilation of L2 vowels to their corresponding categories in L1 (e.g., Flege 1995). The vowel /ɪ/ is spectrally close to Polish /ɨ/ and so our listeners may have categorized /ɪ/ as a vowel different from both /e/ and /ə/. It is not the case for /ə/, which is assimilated to /e/ by speakers of Polish (Bogacka et al. 2006). If this is true, the listeners learnt to use a vowel that sounded different from the two competing vowel categories as an indicator of an unstressed syllable. This pattern has been observed by Zhang and Francis (2010) for Chinese learners of English. Future research should verify if these hypotheses are applicable to listeners on less proficient levels and with a different native language background.

Acknowledgments This research was partly supported by a grant from Polish Ministry of Science and Higher Education (0576/B/H03/2010/38).

References

Anderson, Peggy. J. 1993. *The interstress interval as an indicator of perceived intelligibility among non-native speakers of English*. Doctoral Dissertation, Wichita State University.

Archibald, John. 1992. Transfer of L1 parameter setting: Some empirical evidence from Polish metrics. *Canadian Journal of Linguistics* 37: 301–339.

Beckman, Mary. E. 1986. *Stress and non-stress accent*. Dordrecht: Foris.

Beckman, Mary. E. and Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. In *Laboratory Phonology III: Phonological Structure and Phonetic Form*, ed. P A. Keating, 7–33, Cambridge: Cambridge University Press.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 10: 341–345.

Bogacka, Anna, Geoffrey Schwartz, Paulina Zydorowicz, Monika Połczyńska-Fiszer and Paula Orzechowska. 2006. The production and perception of schwa in second language acquisition: The case of Polish learners of English. In *IFAtuation: A Life in IFA. A Festschrift for Professor Jacek Fisiak on the Occasion of His 70th Birthday*, ed. K. Dziubalska-Kołaczyk, 71–84, Poznań: Wydawnictwo Naukowe UAM.

Braun, Bettina, Kristin Lemhöfer and Anne Cutler. 2008. English word stress as produced by English and Dutch speakers. In *Proceedings of Interspeech 2008*, 1953–1953.

Braun, Bettina, Kristin Lemhöfer and Nivedita Mani. 2011. Perceiving unstressed vowels in foreign-accented English. *Journal of the Acoustical Society of America* 129: 376–387.

Campbell, Nick and Mary E. Beckman. 1997. Stress, prominence, and spectral tilt. In *Proceedings of the ESCA Workshop on Intonation: Theory, Models, and Applications*, eds. A. Botinis, G. Kouroupetroglou, and G. Caraynnis, 67–70, Athens: University of Athens.

Comrie, Bernard. 1967. Irregular stress in Polish and Macedonian. *International Review of Slavic Linguistics* 1: 227–240.

Cooper, Nicole, Anne Cutler and Roger Wales. 2002. Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech* 45: 207–228.

Crosswhite, Katherine. 2003. Spectral tilt as a cue to word stress in Polish, Macedonian, and Bulgarian. In *Proceedings of the XVth International Conference of Phonetic Sciences, Vol. 2*, eds. M. J. Solé, D. Recasens and J. Romero, 767–770, Barcelona: Casual Productions.

Crystal, T. H., House, A. S. 1988. The duration of American English vowels: An overview. Journal of Phonetics 16: 263–284.

Cutler, Anne. 1986. Forbear is a homophone: Lexical prosody does not contain lexical access. *Language and Speech* 29: 201–220.

Cutler, Anne. 2005. Lexical stress. In *The Handbook of Speech Perception*, eds. D. B. Pisoni and R. E. Remez, 264–289, Malden MA: Blackwell Publishing.

Cutler, Anne. 2009. Greater sensitivity to prosodic goodness in non-native than in native listeners. *Journal of the Acoustical society of America* 125: 3522–3525.

Cutler, Anne and Charles. E. Clifton. 1984. The use of prosodic information in word recognition. In *Attention and Performance X*, eds. H. Bouma and D. G. Bouwhuis, 183–196, Hillsdale: Erlbaum.

Cutler, Anne and Dennis. G. Norris. 1988. The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance* 14: 113–121.

Cutler, Anne, Roger Wales, Nicole Cooper and Joris Janssen. 2007. Dutch listeners' use of suprasegmental cues to English stress. In *Proceedings of 16th International Congress of Phonetic Sciences*, eds. J. Trouvain and W. J. Barry, 1913–1916, Saarbrücken.

Dłuska, Maria. 1947. *Prozodia języka polskiego.* Kraków: Polska Akademia Umiejętności.

Dogil, Grzegorz. 1995. Phonetic correlates of word stress. In *Phonetic AIMS*, 1–60, Stuttgart: University of Stuttgart.

Dogil, Grzegorz. 1999. The phonetic manifestation of word stress in Lithuanian, Polish, German and Spanish. In *Word Prosodic Systems in the Languages of Europe*, ed. H. Hulst, 273–311, Berlin: Mouton de Gruyter.

Dupoux, Emmanuel, Sharon Peperkamp and Núria Sebastián-Gallés. 2001. A robust paradigm to study stress "deafness". *Journal of the Acoustical Society of America* 110: 1606–1618.

Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete and Sharon Peperkamp. 2008. Persistent stress "deafness": The case of French learners of Spanish. *Cognition* 106: 682–706.

Dutoit, Thierry, Vincent Pagel, Nicolas Pierret, Ferdinand Bataille and Olivier van der Vrecken. 1996. The MBROLA project: Towards a set of high quality speech synthesizers free of use for noncommercial purposes. In *Proceedings of the Fourth International Conference on Spoken Language Processing*, eds. H. T. Bunnell and W. Idsardi, 1393–1396.

Fear, Beverley D., Anne Cutler, Sally Butterfield. 1995. The strong/weak syllable distinction in English. *Journal of the Acoustical Society of America* 97: 1893–1904.

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-language research*, ed. W Strange, 233–277. Timonium: York Press.

Flege, James E. and Ocke-Schwen Bohn. 1989. An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition* 11: 35–62.

Fokes, Joann and Z. S. Bond. 1989. The vowels of stressed and unstressed syllables in non-native English. *Language Learning* 39: 341–373.

Fokes, J., Z. S. Bond and Mary Steinberg. 1984. Patterns of word stress by native and non-native speakers. In *Proceedings of the 10th International Congress of Phonetic Sciences*, eds. M. Van den Broecke and A. Cohen, 682–686, Dordrecht: Foris.

Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27: 765–768.

Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and Speech* 1: 126–152.

Gay, Thomas. 1978. Physiological and acoustic correlates of perceived stress. *Language and Speech* 21: 347–353.

Gonet, Wiktor, Jolanta Szpyra-Kozłowska and Radosław Święciński. 2010. The acquisition of vowel reduction by Polish students of English. In *Issues in Accents of English 2: Variability and Norm*, ed. E. Waniek-Klimczak, 291–308. Newcastle-Upon-Tyne: Cambridge Scholars Publishers.

Hammond, Robert H. 1986. Error analysis and the natural approach to teaching foreign languages. *Lenguas Modernas* 13: 129–139.

Huss, Volker. 1978. English word stress in the post-nuclear position. *Phonetica* 35: 86–105.

Jassem, Wiktor. 1962. Akcent języka polskiego. *Prace Językoznawcze* 31-1-116.

Jaworski, Sylwester. 2009. Inertial and non-inertial lenition processes. *Poznań Studies in Contemporary Linguistics* 45: 103–129.
Kondo, Yuko. 2000. Production of schwa by Japanese speakers of English. In *Laboratory Phonology V: Acquisition and the Lexicon*, eds. M. B. Broe and J. B. Pierrehumbert, 29–39, Cambridge: Cambridge University Press.
Lee, Borim, Susan G. Guion, and Tetsudo Harada. 2006. Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition* 28: 487–513.
Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge MA: MIT Press.
Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America* 22: 451–454.
Lindblom, Bjorn. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical society of America* 35: 1773–1781.
Mochizuki-Sudo, Michiko and Shigeru Kiritani. 1991. Production and perception of stress-related durational patterns in Japanese learners of English. *Journal of Phonetics* 19: 231–248.
Mol, H. G. and G. M. Uhlenbeck. 1956. The linguistic relevance of intensity in stress. *Lingua* 5: 205–213.
Nowak, Paweł. M. 2006. *Vowel reduction in Polish*. Doctoral Dissertation, University of California.
Ortega-Llebaria, Marta, Maria del Mar Vanrell and Pilar Prieto. 2010. Catalan speakers' perception of word stress in unaccented contexts. *Journal of the Acoustical Society of America* 462–471.
Peperkamp, Sharon and Emmanuel Dupoux. 2002. A typological study of stress "deafness" In *Laboratory Phonology VII*, eds. C. Gussenhoven and N. Warner, 203–240, Berlin: Mouton de Gruyter.
Peperkamp, Sharon, Inga Vendelin and Emmanuel Dupoux. 2010. Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics* 38: 422–430.
Porzuczek, Andrzej. 2008. Stress-dependent syllable duration variability in Polish learners' and native British English pronunciation. In *Issues in Accents of English*, ed. E. Waniek-Klimczak, 64–78, Newcastle-Upon-Tyne: Cambridge Scholars Publishers.
Porzuczek, Andrzej. 2010. The weak forms of TO in the pronunciation of Polish learners of English. In *Issues in Accents of English 2: Variability and Norm*, ed. E. Waniek-Klimczak, 309–324, Newcastle-Upon-Tyne: Cambridge Scholars Publishers.
Rojczyk, Arkadiusz. 2010a. Vowel quality and duration in stressed and unstressed positions in Polish. Poster presented at *GLOW Conference*, April 13, Wrocław Poland.
Rojczyk, Arkadiusz. 2010b. *Temporal and spectral parameters in perception of the voicing contrast in English and Polish*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
Rojczyk, Arkadiusz. 2010c. Production and perception of vowel /æ/ by Polish learners of English. In *Proceedings from the Sixth Symposium on the Acquisition of Second Language Speech*, eds. K. Dziubalska-Kołaczyk, M. Wrembel and M. Kul. [CD ROM Version].
Ramus, Franck, Emmanuel Dupoux and Jacques Mehler. 2003. The physiological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15th International Congress of the Phonetic Sciences*, 337–342.
Slowiaczek, Louisa M. and Daniel A. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 17: 205–212.
Slowiaczek, Louisa M. and Henryka. J. Szymanska. 1989. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 17: 205–212.
Sluijter, Agaath M. C. and Vincent J. van Heuven. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100: 2471–2485.
Sluijter, Agaath M. C., Vincent van Heuven and Jos A. Pacilly. 1997. Spectral balance as a cue in the perception of linguistic stress. *Journal of the Acoustical Society of America* 101: 503–513.
Smoczyńska, Magdalena. 1985. The acquisition of Polish. In *The Crosslinguistic study of Language Acquisition*, ed. D. Slobin, 595–686, Hillsdale, NJ: Lawrence Erlbaum.

Strycharczuk, Patrycja and Peter Jurgec. 2008. Prosodic influences on formant frequencies of Polish vowels. Paper presented at the 3rd Newcastle Postgraduate Conference in Theoretical and Applied Linguistics, July 9, Newcastle.

Turk, Alice E. and James R. Sawusch. 1996. The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America* 126: 367–376.

Tyler, Michael D. and Anne Cutler. 2009. Cross-language differences in cue use for speech segmentation. *Journal of the Acoustical Society of America* 126: 367–376.

Waniek-Klimczak, Ewa. 2002. How to predict the unpredictable: English word stress from a Polish perspective. In *Accents and Speech in Teaching English Phonetics and Phonology*, eds. E. Waniek-Klimczak and P. J. Melia, 221–241, Frankfurt: Peter Lang.

Waniek-Klimczak, Ewa. 2005. *Temporal parameters in second language speech: An applied linguistic phonetics approach*. Lódź: Wydawnictwo Uniwersytetu Łódzkiego.

Weist, Richard, Hanna Wysocka, Katarzyna Witkowska-Stadnik, Ewa Buczkowska and Emilia Konieczna. 1984. The defective tense hypothesis: On the emergence of tense and aspect in child Polish. *Journal of Child Language* 11: 347–374.

Wierzchowska, Bożena. 1980. *Fonetyka i fonologia języka polskiego.* Wrocław: Zakład Narodowy im. Ossolińskich.

Yu, Vickie Y. and Jean E. Andruski. 2010. A cross-language study of perception of lexical stress in English. *Journal of Psycholinguistic Research* 39: 323–344.

Yuan, Jiahong, Stephen Isard and Mark Liberman. 2008. Different roles of pitch and duration in distinguishing word stress in English. In *Proceedings of the 9th Annual Conference of the International Speech Communication Association*, 885, Brisbane, Australia.

Zhang, Yanhong and Alexander Francis. 2010. The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics* 38: 260–271.

Zhang, Yanhong, Shawn L. Nissen and Alexander Francis. 2008. Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America* 123: 4498–4513.

An EMA Study of Articulatory Settings in Polish Speakers of English

Radosław Święciński

Abstract In the present study electromagnetic articulography (EMA) was used to address the question whether articulatory settings are related to foreign-accentedness of speech in Polish speakers of English. The analysis concerns positions of articulators in speech-ready postures assumed shortly before producing an utterance. The objectives of the study are the following: (i) to obtain direct measurement data that could allow for a more detailed description of articulatory settings characteristic of Polish; (ii) to measure and describe articulatory settings in Poles while speaking English (iii) to compare the differences between settings adopted while speaking Polish and English by speakers with near-native English pronunciation and by speakers whose English pronunciation is heavily accented. The conducted experiment confirms the existence of language-specific articulatory settings and their influence on the degree of foreign-accentedness in Polish speakers of English.

1 Introduction

Renewed interest in articulatory settings that can be observed in contemporary linguistic research since the publication of Honikman's article (1964) has resulted in the appearance of a number of empirical studies that make use of the latest technological advances. After a range of works employing acoustic analysis of the speech signal in the description of language-specific habitual postures of the vocal tract (e.g. Laver 1980; Harmegnies and Landercy 1985; Parry et al. 2000; Święciński 2004, 2006), there appeared studies based on more direct measurements

R. Święciński (✉)
Maria Curie-Skłodowska University, Lublin, Poland
e-mail: radek@hektor.umcs.lublin.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_6,

of the configuration of the speech apparatus. These include works founded on x-ray video data (Gick et al. 2004), ultrasound imaging (Wilson 2006; Schaeffler et al. 2008), electropalatography (Schaeffler et al. 2008) or magnetic resonance imaging (MRI) (Ramanarayanan et al. 2010). The present paper also explores the possibility of using latest technology in the description of articulatory settings, namely, with the use of electromagnetic articulogrpahy (EMA).

2 Articulatory Setting

The pronunciation of a language is said to be characterized by certain postures of the speech apparatus which affect the auditory impression of the voice. These postures are generally referred to as the *articulatory setting* and may be defined as pertaining to "general features of pronunciation which underlie all the separate sounds of a language, and which enable us to recognize speakers as 'native' or 'non-native,' irrespective of the clarity of their pronunciation" (Jenner 1997: 38). Thus, articulatory settings are overall articulatory properties shared by all or almost all users of a language. The importance of articulatory settings in pronunciation teaching has been recognized by numerous researchers and a number of, mostly impressionistic, studies were conducted with the aim to explore the applicability of AS teaching in language pedagogy. A thorough review of studies concerning the use of teaching articulatory setting differences in education is presented in Gritter (2010). In relation to the belief that the acquisition of articulatory settings facilitates better performance in the domain of a language's phonetics, dedicated pronunciation practice techniques were developed and tested (Thornbury 1993; Jenner 1997; Wrembel 2001; Święciński 2005; Gritter 2010). Considering the abovementioned theoretical assumptions regarding the importance of articulatory settings in foreign language education, one can infer that Poles who sound like native speakers of English succeeded in disposing of the overall articulatory habits typical of Polish and adopted the setting characteristic of English. This supposition constitutes the basis for the experiment presented in this article.

3 Goals and Scope of the Study

The objectives of the study were the following: (i) to obtain direct measurement data that could allow for a more detailed description of articulatory settings characteristic of Polish; (ii) to measure and describe articulatory settings in Poles while speaking English (iii) to compare the differences between settings adopted while speaking Polish and English by speakers with near-native English pronunciation and by speakers whose English pronunciation is heavily accented, which could verify the claim that articulatory settings are related to foreign-accentedness of speech in Polish speakers of English.

4 The Experiment

In order to accomplish the goals of the study an experiment was designed and conducted. The following sections present its particulars.

4.1 Participants

The experiment was an analysis of articulatory settings in four Polish speakers of English aged between 23 and 30 (2 males and 2 females), whose general level of English proficiency ranged from B2 to C2 according to the Common European Framework of Reference for Languages (Council of Europe 2011). All of the participants claimed to have studied English for at least 8 years. Prior to the recording session they were not informed about the goals of the experiment.

The selection of the participants was performed arbitrarily by the author, the sole criterion being the accentedness of their English pronunciation; two participants displayed a high level of mastery of English phonetics and were accepted for the experiment due to sounding native-like, whereas the pronunciation of the other two could be described as heavily accented with numerous features characteristic of Polish-accented speech. It should also be added that the participants who displayed greater pronunciation skills attended university courses in English phonetics while the other two did not. The speakers were also informally assessed as having neither hearing loss nor speech impediments.

4.2 Experimental Setup

The recording was performed with the use of a three-dimensional AG-500 articulograph that permits digital recording, presentation and evaluation of the movements of the articulators during the production of speech. Its principle function is based on returning spatial coordinates of sensors attached to speech organs. The measurements are performed at 5 ms intervals and are synchronized with simultaneously recorded audio data. Before each recording session, sterilized sensor coils were attached in specific locations on the speaker's body. Reference sensors necessary for calculating head movement correction were placed behind ears and on the bridge of the nose of the recorded participant. Sensors tracing the movements of the lips were glued to the skin just above the upper lip and immediately below the lower lip. (Placing sensors on the very surface of the lips could result in damaging their delicate tissue during sensor removal.) One sensor monitoring the jaw was glued to the gums immediately below the lower incisors. Finally, three sensors were attached to the upper surface of the tongue in locations presented in Fig. 1. Such sensor positioning enabled the inspection of labial, mandibular, tongue tip/blade and tongue body settings.

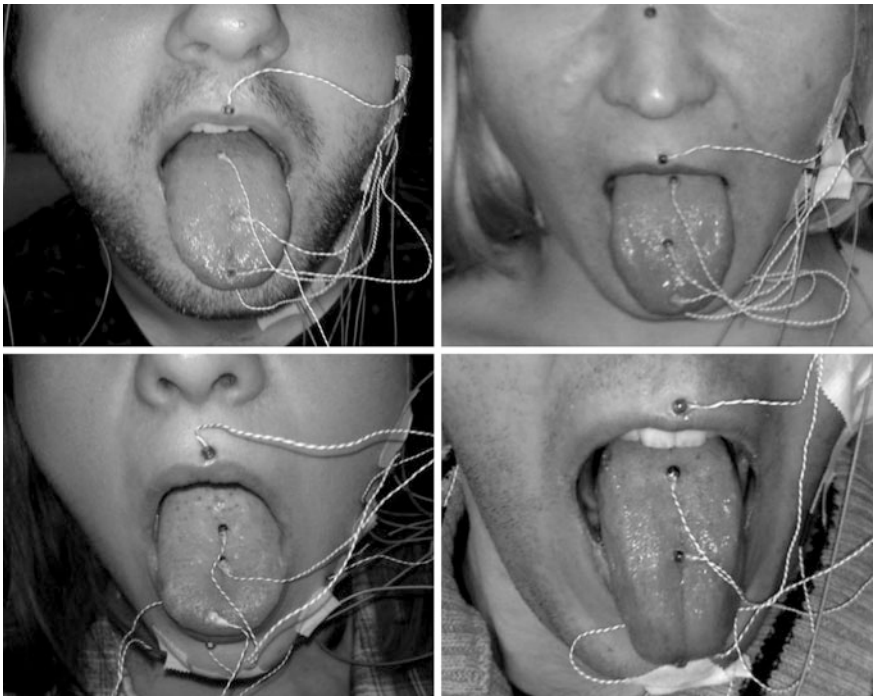


Fig. 1 Sensor placement on the tongues of the participants

4.3 Recorded Material

During the experiment each of the participants was asked to provide answers to ten Polish and ten English general knowledge questions. Usually, the responses were single words. The form of the questions was such that the answer would become obvious only after the last word, e.g. *What is the capital city of Spain?, How much is two plus eight?* Such structuring of the questions ensured that the articulatory movements for the response appeared after the question was presented. Thus, anticipatory positioning of articulators was avoided during stimuli presentation (cf. Hamlet and Stone 1981). Similarly, to minimize any potential anticipatory coarticulation that could affect the pre-speech posture of the vocal tract, the first segments of the desired responses to the stimuli were selected in such a way so as to avoid sounds requiring relatively extensive displacement of the tongue from its neutral position. Hence, the questions presented to the participants elicited words not beginning in interdentals, patals or velars.

The presentation of the stimuli involved both the visual and auditory channel. The audio of the questions was prepared with the use of IVONA text-to-speech synthesizer and embedded in a PowerPoint presentation. Polish and English questions were randomly ordered. Additionally, each auditory prompt was

reinforced by a visual cue for the language involved as the slides which were presented to the participants during the recording displayed words *Question* or *Pytanie* depending on the language of the audio material. For example, the screen displayed text *Question 3* and, subsequently, the participants heard the question *How much is three plus three?*

Following observations by Shaeffler et al. (2008:123) and Wilson (2006:14, 15, 95), the speech prompts were set to appear every 6 s. This timing allowed the speakers to respond to each question, adopt a non-speech rest position between stimuli and minimize the possible carry-over effect of the previously uttered item on the speech-ready posture.

At the beginning of each recording session mid-sagittal outlines of the roofs of the speakers' mouths were traced with an additional sensor. After each recording session, the obtained data were transferred to a PC for position calculation and head-movement correction. The calculations were performed with the use of dedicated software supplied by the producer of AG-500. Thus obtained numerical data consisted of spatial coordinates (in millimeters) revealing the position of particular sensor coils in relation to the centre of the AD500 cube.

4.4 Data Analysis

Articulographic data obtained for each prompt for each participant were plotted separately into graphs showing horizontal and vertical positioning of all relevant sensors against the time axis. Then the plots were manually aligned with corresponding spectrograms of the recorded audio signal as presented in Fig. 2. The next step involved identifying speech-ready postures (SRPs).

4.5 Speech-Ready Postures

Contemporary research into articulatory settings of languages focuses on describing the configuration of the vocal tract in a form which is least affected by speech sound segments. Thus, a number of works (e.g. Gick et al. 2004; Wilson 2006; Shaeffler et al. 2008) follow the assumption, also adopted in this study, that, before making an utterance, human speech organs adopt a language-specific preparatory posture which reflects the language's articulatory setting. Hence, most AS-centred research focuses not on describing the vocal tract during speech production, but rather on the investigation of the speech apparatus shortly before articulation. Using this approach, it is possible to decrease the coarticulatory effect of the segmental context of an utterance on the articulatory setting. Gick et al. (2004: 220) claim that such speech rest positions (inter-utterance rest positions) are specified in a manner similar to actual speech targets. Authors have applied numerous names to this preparatory posture, e.g. inter-speech posture/inter-utterance position (ISP) (2004; Wilson 2006;

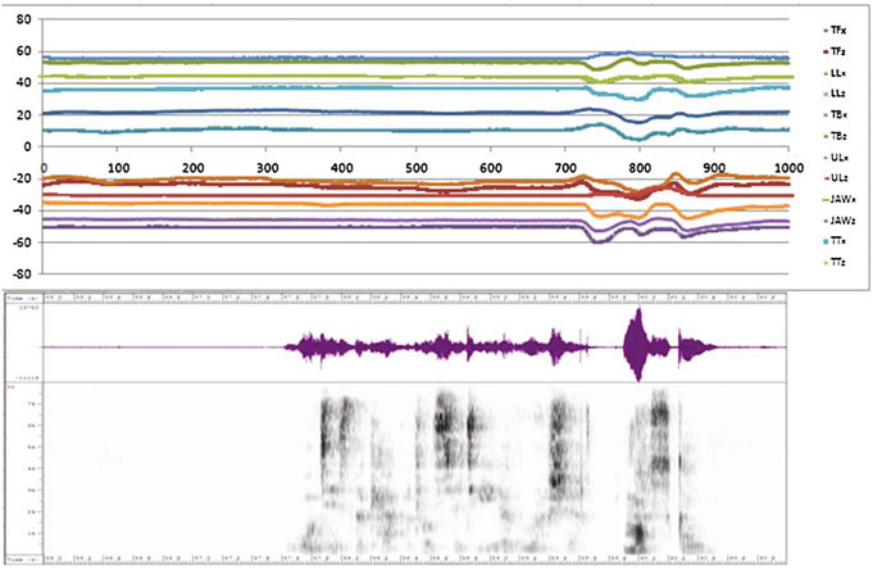


Fig. 2 Sensor movement trajectories aligned with the visualisation of the prompt and response (Speaker 1). The abbreviations are explained in the *Results* section

Shaeffler et al. 2008); basic speech posture (Öhman 1967); speech-ready position (SRP) (Ramanarayanan et al. 2010). For a more detailed discussion of speech rest positions cf. Wilson 2006. On semantic grounds, the author chose to use the last term as it best reflected the design of the experiment. The settings scrutinized here were not those that appear within longer utterances but before short instances of speech activity. It should also be mentioned that the language-specific SRP ought to be distinguished from the absolute resting vocal tract position which is characterised by an almost closed vocal tract with a relatively small jaw angle and a narrow pharynx (Ramanarayanan et al. 2010).

During data preparation for the present study, an SRP was identified in the case of presence of any notable change in the overall positioning pattern between the time when the prompt appeared and the first speech gesture. It was accepted if the calculated movement trajectories indicated a motion towards some configuration, followed by smooth movement away from it towards the first segment. In Fig. 2 one can notice that at the end of the prompting question all sensors begin to move and, after reaching certain values, most of the articulators remain in the adopted positions for some time; following that, the speech organs begin to move towards the articulatory targets for the first segment of the response. The SRP in this token is marked with a grey frame.

It is worth mentioning that the speakers participating in the experiment appeared to adopt different strategies in realizing their SRP's. As seen in Fig. 2, during the presentation of the prompt, the participant held her mouth closed with

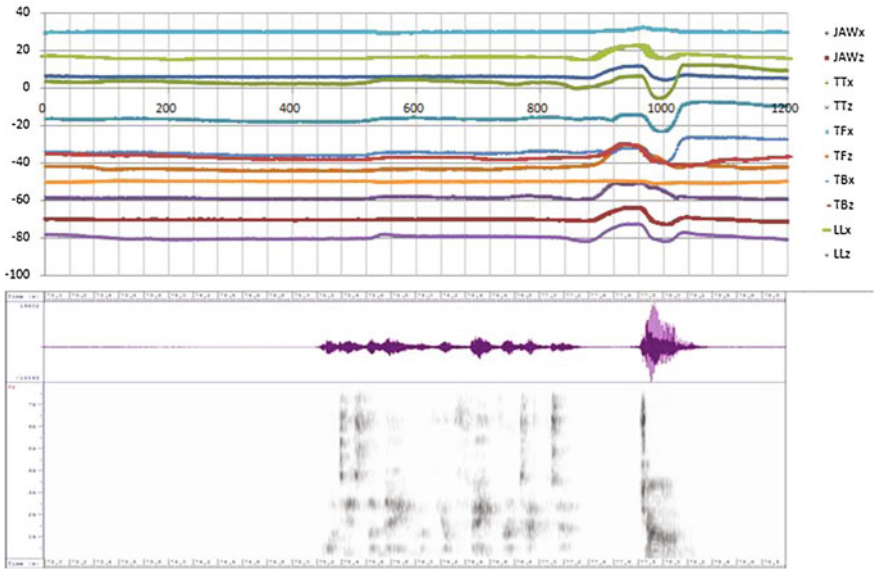


Fig. 3 Sensor movement trajectories aligned with the visualisation of the prompt and response (Speaker 2)

the tongue touching the roof of the mouth. She would lower the jaw and position the speech organs in a speech-ready posture just before the end of the question. Another approach can be observed in Fig. 3. Here the participant prepared for speaking in advance, shortly after identifying the language to be used on hearing the initial part of the prompt.

5 Results

Following the identification of speech-ready postures, a sample of articulographic data from each SRP was selected for each token. Four tokens were excluded from analysis as it was difficult to distinguish SRPs in them. Samples were selected in such locations on the plots where most of the articulators appeared to be stable after reaching their speech-ready posture targets. Then positional data of sensors were extracted and analysed for each speaker and language. Tables 1, 2, 3, 4 show mean positional values of sensor coils in the horizontal (x) and vertical (z) dimension. Particular sensors are labeled as follows: UL = upper lip, LL = lower lip, TT = tongue tip, TF = tongue front, TB = tongue back. The results of the comparison reveal statistically significant differences ($p = 0.05$) between samples related to tongue position prior to speaking English and Polish in Speakers 1 and 2. These were the participants who displayed near-native pronunciation. In the case

of Speaker 1, the central line of the tongue was on average 2 mm more to the front for English SRP's than for Polish. Moreover, the front and back of the tongue were elevated by approximately 2 mm more for English compared to Polish speech-ready postures. Speaker 2 displayed similar results with statistically significant difference in tongue vertical positioning between English and Polish SRP's. No statistically valid differences were attested for Speakers 3 and 4 at the same level of probability. Tables of results can be found in Appendix A.

6 Conclusions

Statistically significant differences that were observed in SRPs of the tongue assumed before speaking Polish and English in Speakers 1 and 2 and lack of such differences in Speakers 3 and 4 support the claim that articulatory settings are related to foreign-accentedness. The participants who displayed a greater command of English phonetics (1 and 2) appear to have learnt to vary their habitual posture of the vocal tract depending on the language they are going to speak, which, most probably, contributes to their pronunciation being perceived as native-like. The present study is in line with the experimental findings of Gick et al. (2004) and offers further evidence for the existence of articulatory settings as such.

All of the participants positioned the lips and the jaw in a similar manner for both languages. The results cannot be said to support a claim that mandibular and labial settings are the same in Polish and English since the phonetically trained participants, even though displaying a high degree of proficiency in English phonetics, may not have acquired all facets of English pronunciation.

The data show that during SRP's in the case of both languages, the tongue body moves farther front and higher than in the rest position.

The participants who displayed significant differences in their SRP's for Polish and English adopted the strategy of bringing the tongue forward in the midsaggital plane when preparing for English pronunciations. This result may have two interpretations. Firstly, it can be presumed that the whole mass of the tongue moved to the front before speaking English. Secondly, the data could reflect a habitual groove along the centre of the tongue in Polish with the sides of the tongue remaining in a similar position as for English. This issue is still to be verified.

Appendix A

Table 1 Mean sensor positions during SRPs in Speaker 1 (F1)(in millimeters)

	ULx	ULz	LLx	LLz	TTx	TTz	TFx	TFz	TBx	TBz	JAWx	JAWz
Polish												
Mean	56.7	−28.9	49.6	−57.2	32.8	−42.3	22.2	−26.8	12.1	−21.5	41.7	−51.1
SD	1.58	0.69	0.74	1.89	0.73	1.53	1.54	1.50	1.85	1.70	0.77	1.65
English												
Mean	56.8	−29.3	49.9	−56.4	33.9	−41.5	24.2	−25.6	14.3	−19.5	41.7	−50.8
SD	1.46	1.22	1.04	2.36	1.26	2.47	1.01	0.92	1.11	1.06	0.68	1.41

Shaded columns show statistically significant differences between the values for Polish and English at $p = 0.05$

Table 2 Mean sensor positions during SRPs in Speaker 2 (M1) (in millimeters)

	ULx	ULz	LLx	LLz	TTx	TTz	TFx	TFz	TBx	TBz	JAWx	JAWz
Polish												
Mean	29.7	−49.5	17.6	−77.9	1.9	−55.6	−17.6	−42.6	−35.3	−39.0	7.1	−69.6
SD	0.39	0.49	1.69	2.48	3.36	3.39	2.58	2.32	2.25	1.40	1.56	1.96
English												
Mean	29.6	−49.6	17.7	−77.4	5.5	−56.9	−15.0	−43.3	−33.3	−38.5	7.3	−69.0
SD	0.36	0.30	1.47	1.96	2.45	1.13	1.84	1.69	1.45	1.55	1.25	1.46

Shaded columns show statistically significant differences between the values for Polish and English at $p = 0.05$

Table 3 Mean sensor positions during SRPs in Speaker 3 (in millimeters)

	ULx	ULz	LLx	LLz	TTx	TTz	TFx	TFz	TBx	TBz	JAWx	JAWz
Polish												
Mean	40.9	−100.5	33.1	−113.7	20.1	−98.3	7.7	−82.4	−4.2	−75.4	24.0	−109.8
SD	3.97	4.34	3.70	5.77	2.95	5.71	4.04	6.70	3.26	6.82	3.44	4.56
English												
Mean	42.2	−101.6	34.4	−115.3	20.1	−99.0	6.4	−84.8	−4.8	−78.1	25.0	−110.6
SD	3.35	3.68	4.70	2.34	4.63	3.38	5.12	4.66	3.83	5.71	4.13	2.47

Table 4 Mean sensor positions during SRPs in Speaker 4 (in millimeters)

	ULx	ULz	LLx	LLz	TTx	TTz	TFx	TFz	TBx	TBz	JAWx	JAWz
Polish												
Mean	75.5	3.3	71.5	−17.4	49.2	−5.7	34.9	1.8	12.2	−4.2	57.9	−18.4
SD	0.34	0.52	0.61	3.41	1.02	2.99	1.69	2.89	0.86	4.07	0.66	2.16
English												
Mean	75.4	3.4	71.4	−19.6	48.9	−7.8	34.3	0.3	12.5	−6.3	57.6	−20.0
SD	0.47	0.50	0.67	3.78	1.52	4.33	2.06	3.69	1.15	3.72	0.88	2.84

References

Council of Europe. 2011. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Council of Europe.

Gick, Bryan, Ian Wilson, Kenneth Koch and Clare Cook. 2004. Language-specific articulatory settings: Evidence from inter-utterance rest position. *Phonetica*, 61, 220-233.

Gritter, Eliene. 2010. *Repeat after me: Mississippi, Mississippi, Mississippi. A study into the use of teaching articulatory setting differences in secondary education*. MA thesis: Rijksuniversiteit Groningen.

Hamlet, Sandra, L., Maureen L. Stone. 1981. "Pre-speech posturing of the mandible in relation to the jaw activity during speech," *J. Phonetics*, 9, 425-436.

Harmegnies, Bernard, Albert Landercy. 1985. Language features in the long-term average spectrum. *Rev. Phonét. appl.* 73–75, 69–80.

Honikman, Beatrice. 1964. Articulatory settings. In Abercrombie, David, Dennis B. Fry, Peter A.D. MacCarthy, Norman C. Scott, and John L.M. Trim (eds.). 1964. In honour of Daniel Jones: Papers contributed on the occasion of his eightieth birthday 12 September 1961. London: Longmans., pp. 73-84

Laver, John. 1980. *The Phonetic Description of Voice Quality,* Cambridge: Cambridge University Press.

Jenner, Bryan. 1997. "The English Voice", *Approaches to Pronunciation Teaching,* Brown, Adam [ed.], Hemel Hempstead: Prentice Hall International: 38-46.

Öhman, Sven, E. G. 1967. Peripheral Motor Commands in Labial Articulation. *Speech Transmission Laboratory – Quarterly Progress and Status Report (Royal Institute of Technology (KTH), Stockholm)*. 4/1967, 30-63.

Parry, John, J., Ian S.Burnett, Joe F. Chicharo. 2000. Language-specific phonetic structure and the quantisation of the spectral envelope of speech. *Speech Commun*. 32, 229–250.

Ramanarayanan, Vikram, Dani Byrd, Louis Goldstein, Shrikanth Narayanan. 2010. "Investigating articulatory setting – pauses, ready position, and rest – using real-time MRI," *Proceedings of Interspeech 2010*, Makuhari, Japan.

Schaeffler, Sonja, James Scobbie, Ineke Mennen. 2008. An evaluation of inter-speech postures for the study of language-specific articulatory settings. In Rudolph Sock, Susanne Fuchs & Yves Laprie (eds.), *Proceedings of the 8th International Speech Production Seminar*, Strasbourg, France: INRIA, 121-124.

Święciński, Radosław. 2004. 'Articulatory Setting in Polish and its Implications for Teaching English Pronunciation to Poles.' In: *Zeszyt Naukowy Instytutu Neofilologii (3), Zeszyty Naukowe PWSZ w Koninie.* 1/2004 (4). 141-150.

Święciński, Radosław. 2006. 'Teaching English articulatory setting features to Polish students of English – a study of phonation.' In Sobkowiak, Włodzimierz & Ewa Waniek-Klimczak (red.) *Dydaktyka fonetyki języka obcego w Polsce.* Konin: Wydawnictwo PWSZ w Koninie. 203-215.

Thornbury, Scott. 1993. Having a good jaw: Voice-setting phonology. *ELT Journal*, *47*, 126-131.

Wilson, Ian. (2006). Articulatory settings of French and English monolingual and bilingual speakers. Ph.D Thesis. University of British Columbia.

Wrembel, Magdalena. 2001. Innovative approaches to the teaching of practical phonetics. In *Proceedings of the Teaching and Learning conference PTLC2001*, London: UCL, 63-66.

Part II
Pedagogy

The Effect of Explicit and Implicit Corrective Feedback on Eliminating Pronunciation Errors

Mirosław Pawlak

Abstract Recent years have witnessed an abundance of research on the effectiveness of different forms of corrective feedback on the acquisition of a variety of grammatical features (e.g. Russell and Spada in Synthesizing research on language learning and teaching. John Benjamins, Amsterdam, 2006; Ellis in Studies in second language acquisition 32:335–349, 2010; Li in Lang Learn 60:309–365, 2010; Lyster and Saito in Stud Second Lang Acquis 32:265–302, 2010; Pawlak in Heteronomie glottodydaktyki: Domeny, pogranicza i specjalizacje nauczania języków obcych. Instytut Filologii Angielskiej, Poznań, 2010; Sheen and Ellis in Handbook of research in second language learning and teaching. Routledge, London and New York, 2011; Pawlak in Error correction in the foreign language classroom: reconsidering the issues. Adam Mickiewicz University and State School of Higher Professional Education in Konin Press, Poznań–Kalisz–Konin, 2012), which demonstrated, that the treatment of errors in the course of communicative activities results in increased control over the targeted linguistic features, not only in terms of explicit but also implicit knowledge, with the qualification that such pedagogic intervention should be focused, intensive and consistent (cf. Leeman in Practice in a second language: perspectives from applied linguistics and cognitive psychology. Cambridge University Press, Cambridge, 2007; Ellis in L2 J 1:3–18, 2009; Pawlak in Heteronomie glottodydaktyki: Domeny, pogranicza i specjalizacje nauczania języków obcych. Instytut Filologii Angielskiej, Poznań, 2010; Sheen and Ellis in Handbook of research in second language learning and teaching. Routledge, London and New York, 2011; Pawlak in Error correction in the foreign language classroom: reconsidering the issues. Adam Mickiewicz University and State School of Higher Professional Education in Konin Press, Poznań–Kalisz–Konin, 2012). Little research specifically addresses the impact of

M. Pawlak (✉)
Adam Mickiewicz University, Kalisz, Poland
e-mail: pawlakmi@amu.edu.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_7, © Springer-Verlag Berlin Heidelberg 2013

various techniques of providing corrective feedback on the acquisition of foreign language pronunciation. The paper reports the findings of a pilot study which compared the effect of explicit (overt) and implicit (covert) correction on eliminating pronunciation errors committed by advanced learners of English. The study involved 36 English philology students and took the form of a quasi-experiment with two experimental and one control groups. The students in the experimental groups took part in communicative activities in the course of which their mispronunciations of selected words were corrected explicitly and implicitly while the control group focused on other tasks. Ability to pronounce the targeted words was measured on a pretest and a posttest, both of which included a reading text and a free production task. The results serve as a basis for a handful of recommendations on how pronunciation errors should most beneficially be dealt with and some suggestions for further research.

1 Introduction

Although the beginnings of research into the effectiveness of oral corrective feedback can be traced back to the 1970s and 1980s, as evidenced, for example, by the studies conducted by Ramirez and Stromquist (1979), Brock et al. (1986) or Chaudron (1986), it is fair to say that empirical investigations in this area gained momentum with the emergence of the concept of *focus on form* (Long 1991) and the resultant revival of interest in form-focused instruction (FFI). Although the ways in which focus on form has been conceptualized and operationalized for the purposes of research projects have undergone major modifications over the last two decades (cf. Williams 2005; Pawlak 2007; Loewen 2011), the pedagogical intervention it recommends is based on the assumption that learners' attention should be directed at linguistic features as they are engaged in meaning and message conveyance, as is the case with the performance of communicative tasks. As a result, it differs in a fundamental way from both a *focus on meaning*, which by and large rejects all variants of formal instruction and error correction, and a *focus on forms*, which entails careful preselection of the structures to be taught as well as reliance on the presentation—practice—production (PPP) procedure, in many cases with little concern for the communication of real meanings. The most obvious manner in which focus on form can be accomplished is through the provision of corrective feedback (CF) in the course of communicative activities, with the effect that such error correction should be viewed as a key option in form-focused instruction.

Support for the treatment of learners' errors as they are trying to attain genuine communicative goals can be derived from a number of theoretical positions, both psycholinguistic and sociolinguistic in nature, such as the Noticing Hypothesis (Schmidt 1990, 2001), the revised Interaction Hypothesis (Long 1996), the Output Hypothesis (Swain 1995, 2005), Skill-Learning Theory (DeKeyser 1998, 2007),

the Counterbalance Hypothesis (Lyster and Mori 2006), Sociocultural Theory (Lantolf 2006; Lantolf and Thorne 2007), and Relevance Theory (Niżegorodcew 2007). There is also abundant empirical evidence demonstrating that correction provided during meaning-based activities contributes to enhanced control over specific language forms not only in terms of explicit, declarative knowledge, but also, much more importantly, implicit, procedural knowledge. For these benefits to accrue, however, such pedagogic intervention should be *focused* (i.e. it should be directed at a single feature or a limited set of such features), *intensive* (i.e. the item is targeted many times), and *consistent* (i.e. the intervention should target all the errors in the use of the structure and be prolonged) (Russell and Spada 2006; Leeman 2007; Li 2010; Lyster and Saito 2010; Ellis 2009, 2010; Sheen 2010; Pawlak 2010; Sheen and Ellis 2011; Pawlak 2012). Moreover, questions remain as to the instructional target, the value of input-providing and output-inducing corrective feedback, the utility of explicit and implicit correction, the effects of priming, or the impact of individual, contextual and linguistic variables (cf. Ellis 2010; Sheen and Ellis 2011; Pawlak 2012). Another problem is that much of the research conducted to date has focused on the impact of CF on the acquisition of grammatical features, with studies of other target language subsystems, such as pronunciation, being few and far between.

The present paper is intended to address the unfortunate gap in the existing research by reporting the findings of a study which set out to compare the value of explicit (overt) and implicit (covert) corrective feedback in eliminating persistent pronunciation word-level errors in the speech of advanced Polish learners of English. At the very outset, the choices that practitioners have at their disposal when responding to pronunciation errors will be outlined and the findings of the scant research in this area will briefly be considered. This will be followed by an overview of empirical investigations that have not focused on pronunciation as such but are relevant to the present discussion since they have investigated the effectiveness of explicit and implicit error correction in the acquisition of specific linguistic features. The second part of the paper will be devoted to the presentation of the methodology of the study conducted by the present author and the discussion of its results, which will serve as a basis for tentative pedagogic recommendations as well as directions for future research into corrective feedback addressing pronunciation errors.

2 Corrective Feedback and Pronunciation Instruction

As mentioned above, current research into the role of corrective feedback has hardly been extended into pronunciation instruction, with the effect that many of the assumptions made in this respect have to be viewed as somewhat speculative, as they are not supported by ample empirical evidence. One attempt to discuss the methodological options in the correction of pronunciation errors has been made by Pawlak (2004), who, drawing on the available literature devoted to form-focused

instruction and corrective feedback provided in response to inaccurate uses of grammar structures (e.g. Allwright and Bailey 1991; Doughty and Williams 1998; Ellis 2001, 2006), identified the following decisions which have to be considered by teachers and researchers:

(1) *planned* vs. *incidental* correction, a distinction which is related to whether the target form is preselected in advance or the pedagogic intervention takes place on the spur of the moment; in the former case, the CF is *focused* and *intensive* because a specific aspect of pronunciation is addressed on multiple occasions, while in the latter it is *unfocused* and *extensive* on account of the fact that multiple pronunciation features are dealt with, many of them only once;
(2) *selecting features to be corrected*, which is connected to some extent with the planned—incidental distinction made above, but, particularly in the case of pronunciation, it can also determine whether the intervention is intended to trigger *item-learning* (i.e. learning separate, discrete instances) or *system-learning* (i.e. learning the rules underlying the occurrence of such instances); thus, when correction is directed at all the inaccurate uses of a particular segment, including all allophonic variation, the expected outcome is system-learning, whereas, CF focusing on the pronunciation of a set of problematic words, such as that employed in the study reported below, is mainly aimed to promote item-learning;
(3) *timing of corrective feedback*, which is related to whether the intervention immediately follows the commission of the error, thus being *immediate*, or it is deferred until the end of a learner's contribution, the completion of an activity, or even the following lesson, in which case it is *delayed*;[1]
(4) *choosing a corrective technique*, with the most influential distinctions being made between *input-providing* and *output-promoting* corrective feedback, on the one hand, and *explicit (overt)* and *implicit (covert)* error correction on the other; the former is related to the requirement for output modification and self-correction and it is typically operationalized as the use of *recasts* (i.e. reformulations of a learner's utterance that do not have to be incorporated by him or her) and *prompts* (e.g. clarifications requests or elicitations that put the onus on the learner to fix the problem); the latter, which is of particular relevance to the present paper, has to do with the level of awareness on the part of the learner that the teacher's reaction is corrective in nature, and it is most frequently operationalized as the provision of *recasts* or *clarification requests* as opposed to *direct correction* or *elicitation*, usually accompanied by some kind of *metalinguistic information*;

[1] In fact, it is possible to make a finer-grained distinction between *delayed correction*, which occurs within the confines of the same lesson, preferably after the completion of a turn or a speech, and *postponed correction*, which may take place in the following lesson and deal with a range of errors committed by learners (cf. Allwright and Bailey 1991; Pawlak 2012). Such a differentiation, however, is of no relevance to the present discussion.

(5) *determining the source of error correction*, which is reflective of whether the negative feedback comes from *the teacher*, which is the most common in the vast majority of classrooms, *the learner who has erred*, in which case we are dealing with *self-correction*, or another student, when *peer-correction* takes place.

Two important caveats are in order at this point. First, it should be emphasized that the decisions made in the five areas are intricately intertwined, because, for example, the corrective techniques included in (4) can only be used when errors are corrected immediately (3), and the occurrence of self-correction (5) hinges to a large extent upon reliance on output-promoting corrective techniques (3). Second, the decisions made with respect to the dimensions mentioned above are to a large extent a function of the type of activity in which the correction of pronunciation errors takes place. To be more specific, the degree of planning, the choice of the instructional target, the timing, manner and source of correction are likely to vary depending on whether CF is provided during accuracy-based activities, such as reading aloud or minimal pair practice, which is likely to contribute to the growth of explicit knowledge, or fluency-oriented ones, such as focused communication tasks (i.e. such requiring the use of a specific linguistic feature for their successful completion), in which case the intervention may have a somewhat more direct impact on the development of implicit knowledge (cf. Pawlak 2006, 2012).

There is a paucity of research focusing on the value of corrective feedback in teaching pronunciation, with the effect that there is little empirical support for the superiority of any of the methodological options discussed above or constellations of such options. Some insights in this area can be gleaned from descriptive studies of error correction, such as those conducted by Lyster and Ranta (1997), Lyster (1998) or Panova and Lyster (2002), in which pronunciation problems constituted one of the categories of errors considered in terms of uptake and repair. All of these studies demonstrated that although recasts, which are input-providing and relatively implicit, were the most frequent, they were the least likely to generate output modifications, much less so than different forms of negotiation of form, or prompts. They proved to be the most effective in the case of pronunciation errors, but only because these were corrected during reading-aloud activities, in which their corrective function was transparent to the learner, and it is doubtful that they would have been responded to in the same way during fluency-oriented activities. In the Polish context, the role of error correction in pronunciation instruction has been investigated by Szpyra-Kozłowska et al. (2004), Jelska-Cydzik (2006), and Pawlak and Pospieszyńska (2003). Szpyra-Kozłowska et al. (2004) looked into the provision of corrective feedback during accuracy-based activities and conducted two experiments, the first of which compared the effectiveness of the correction of selected mispronunciations and all the major errors, or, to use the terminology introduced in (1) above, focused and unfocused CF, and the other explored the value of pronunciation instruction with and without error treatment. The found that: (1) extensive, unfocused correction was more effective than intensive focused feedback, and (2) instruction augmented with the provision of CF was superior to

instruction without it since, as they comment, "(…) applying systematic correction results in the improvement of English pronunciation while its absence slows down phonetic progress" (2004, p. 131) Jelska-Cydzik (2006), in turn, looked into the role of explicit instruction, which included a phase involving self-correction and peer-correction, in helping Polish learners of English gain greater control over word stress, thus focusing on item-learning rather than system-learning. The intervention proved to be effective in raising the learners' awareness of the stress patterns of particular words, but the weakness of the study is that the participants were requested to mark the primary stress in 40 instructional targets rather than to produce them. The influence of corrective feedback on the occurrence of item-learning was also explored by Pawlak and Pospieszyńska (2003), who investigated the contribution of implicit correction in the form of recasts provided in communicative activities on 32 secondary school learners' ability to pronounce 36 problematic words. Even though the scores on the pretest, immediate and delayed posttest demonstrated a dramatic improvement, these findings have to be regarded with circumspection in view of the fact that there was no control group or a comparison group that would have received more explicit correction. Finally, worth mentioning is a recent study carried out by Saito and Lyster (2011), who compared the effects of form-focused instruction with and without corrective feedback in the form of recasts on the acquisition of the English /ɹ/ by 65 Japanese learners. They conducted acoustic analysis of the frequency values of the third formant in controlled and spontaneous speech production on a pretest and a posttest, and found that that FFI with CF was more effective than FFI alone or a sheer focus on meaning.

3 Research into Explicit and Implicit Corrective Feedback

In view of the fact that no studies conducted to date have compared the effects of explicit and implicit correction on the acquisition of pronunciation features, it seems warranted to briefly consider the findings of research that has considered the effectiveness of these CF options in teaching other language subsystems. Most of these empirical investigations provide rather clear-cut evidence for the greater value of corrective techniques that are explicit and therefore can unambiguously be interpreted by learners as providing negative evidence rather than constituting a mere reaction to the content of their utterances. This is evident in Sheen's (2010, p. 173) comment that "(…) it would appear that explicit CF is more effective than implicit recasts", as well as Sheen and Ellis's (2011, p. 607) assessment that "(…) explicit feedback in conjunction with metalinguistic clues is more likely to result in learning than recasts". A crucial caveat, however, is that this advantage is much more evident in real classrooms than in laboratory settings, which can be ascribed to the fact that in the latter case the correction can hardly be misinterpreted, the

nature of one-on-one interactions makes the participants more likely to attend to recasts, and it is more feasible to provide continuous treatment on single target language features (Spada and Lightbown 2009; Li 2010; Lyster and Saito 2010).

Support for the greater benefits of more explicit feedback types comes, among others, from the studies conducted by Carroll and Swain (1993), Rosa and Leow (2004), Ellis et al. (2006), Ellis (2007), Sheen (2007), as well as Varnosfadrani and Basturkmen (2009). Carroll and Swain (1993) investigated the acquisition of English dative verbs by adult native speakers of Spanish performing mechanical exercises under five conditions: (1) direct metalinguistic feedback (i.e. information about the correct form), (2) explicit rejection (i.e. information about the occurrence of an error), (3) recasts, (4) indirect metalinguistic feedback (i.e. a question about the correctness of the utterance), and (5) no feedback. Although all the types of correction resulted in improvement on recall production tasks, the group receiving direct metacognitive CF outperformed the others. The computer-based study undertaken by Rosa and Leow (2004) focused on the acquisition of contrary-to-the-fact conditional in Spanish in the context of performing focused communication tasks, requiring the use of the targeted feature for their successful completion (cf. Ellis 2003). The group that was provided with explicit feedback on accurate and inaccurate utterances as well as the opportunity to self-correct scored significantly higher on multiple-choice recognition tests and written controlled production tests than the one in which implicit correction was employed, although the students in this group still did better than the controls. Elli, Loewen and Erlam (2006) conducted a classroom-based study which investigated the impact of explicit and implicit CF on the acquisition of the regular past tense '-ed' ending by lower-intermediate adult ESL learners in New Zealand. Also here the feedback groups did better than the controls, and prompts involving the use of metalinguistic information proved to be more effective than recasts in the long run, particularly with respect to the development of implicit knowledge as well as the occurrence of system-learning. Similar in design was the study undertaken by Ellis (2007), which focused on the acquisition of the '-ed' ending and the comparative, and also corroborated the superiority of more overt corrective feedback options. Analogous results have been reported by Sheen (2007), who compared the effects of direct correction coupled with metalinguistic explanation and recasts on the acquisition of English articles by intermediate-level students, as well as Varnosfadrani and Basturkmen (2009), who found that explicit correction resulted in better scores on individualized post-tests, with the caveat that its contribution was mediated by the participants' developmental readiness (i.e. it benefitted developmentally early features more than developmentally late features).

This overview would surely be incomplete without mentioning the research projects which have failed to provide clear evidence for the greater effectiveness of explicit CF options, such as those conducted by Kim and Mathes (2001), Sanz (2003), and Loewen and Nabei (2007). The first of these was a replication of the study carried out by Carroll and Swain (1993), which was considered above, but it failed to identify statistically significant differences between any of the feedback conditions. Sanz (2003), in turn, examined the provision of computer-mediated

corrective feedback but also here the group that received explicit metalinguistic feedback performed similarly to the one that had the benefit of implicit correction (i.e. an invitation to try again). Finally, Loewen and Nabei (2007) compared the impact of CF in the form of metalinguistic information, recasts and clarification requests on the acquisition of question formation by adult Japanese learners, and reported similar gains in all the three conditions, with no such improvement in the control group. A possible explanation for the fact that these studies have not yielded evidence for the greater value of explicit feedback may be the inclusion of a small number of subjects receiving a variety of instructional treatments (Kim and Mathes 2001), the absence of grammatical explanation and a comprehension-based nature of the treatment in a laboratory setting (Sanz 2003), or the use of an additional category of output-prompting feedback (Loewen and Nabei 2007). Nevertheless, it should be emphasized that none of these research projects have demonstrated the superiority of implicit error correction, which indicates that such feedback can at best work as well as explicit feedback, on condition that favorable conditions are in place.

4 Research Questions and the Design of the Study

The research project reported in the present paper constitutes part of a larger-scale empirical investigation of the role of CF in teaching pronunciation, much more complex in design, and thus it should be regarded as a pilot study. It aimed to investigate the effect of explicit and implicit corrective feedback, consistently delivered in the course of communicative tasks, on eliminating persistent word-level pronunciation errors. More specifically, the study addressed the following research questions:

- Does correction aid students in eliminating persistent errors?
- Does correction affect both explicit and implicit knowledge?
- Is there a difference in the effect of explicit and implicit corrective feedback with respect to the two types of knowledge?
- What variables may affect the impact of corrective feedback?

The subjects were 36 students enrolled in the 2nd year of a 3-year BA program in the Department of English Studies, who had received intensive instruction in English which included courses in grammar, speaking, writing, integrated skills and pronunciation. Most of the subjects had considerable experience in learning English prior to their enrollment in the BA program and their proficiency level could be described as B1 in terms of the scales included in the *Common European Framework for Languages*, with the qualification that there was much individual variation. The participants attended three intact classes (i.e. such into which the students had been grouped at the beginning of the BA program) that were comparable in all important respects (i.e. level, length of study, motivation), and were randomly designated as two experimental groups, one receiving implicit (IG, $N = 14$) and the other explicit feedback (EG, $N = 12$), and a control group (CG, $N = 10$).

The pedagogical intervention was intended to trigger *item-learning* rather than *system-learning* (see Sect. 2), since the instructional targets were words that were known to represent a major learning challenge to the students in terms of their pronunciation rather than systematic segmental or suprasegmental features. The justification for this approach can be found in Szpyra-Kozłowska and Stasiak (2010, p. 164), who make the point that "eliminating phonologically deviant words (…) should be the primary concern of language instructors", adding that phonetic instruction confined to segments and suprasegments is highly unlikely to eradicate such mispronunciations. While it is not argued here that intervention of this kind should entirely replace the introduction and practice of more global pronunciation features, it is warranted to assume that a focus on word-level phonetic errors can be beneficial and it should therefore complement the system learning and teaching that is the staple of pronunciation courses in English departments. In line with this reasoning, 90 problematic words were identified for the participants of the study on the basis of observations of practical English and content classes as well as the suggestions made by regular teachers of the students taking part in the experiment (see the highlighted words in the text in Appendix 1). Consequently, apart from verifying the claims derived from theory and research, the instruction in the form of error correction was motivated by learner need, so the study can be said to possess high ecological validity.

As can be seen from Fig. 1, which diagrammatically presents the design of the research project, it spanned the period of 4 weeks and involved the use of a pretest–posttest procedure. While the students in the CG followed their regular activities, those in the two experimental groups had the benefit of an instructional treatment that was massed over the period of 1 week and took four sessions, 45–50 min each, held in the course of their practical English classes. On each of the four occasions, they were asked to perform communicative tasks (decision-making, information-gap, etc.) calling for the use of the 90 problematic words, with these activities first being done in pairs or small groups and the outcomes later being reported to the whole class. All the errors involving the targeted words were corrected during group-work and whole-class interaction, while all the other inaccuracies, irrespective of whether they were related to pronunciation, grammar, lexis or pragmatics, were ignored, which makes the intervention planned, focused and intensive. The difference between the IG and EG was that in the former the CF was implicit and took the form of *recasts* and *clarification requests*, and in the latter the correction was explicit and involved the use of *direct error correction* or *elicitation*, sometimes accompanied by *metalinguistic cues*. Thus, when a student mispronounced a targeted word as in: *Frank went to the hospital in July /dʒuli/*, (1) in the IG the corrective reaction would have been: *He went to the hospital in JULY OK, and what happened later?* (a recast with the targeted word receiving additional stress) or *Could you repeat? What did you say?* (a clarification request requiring self-correction), and (2) in the EG, it would have been: *We say JULY, remember?* (direct correction with emphatic stress), or *Actually there is a problem because you need a diphthong* (elicitation plus a metalinguistic cue). It should also be mentioned that all the teachers of the three groups were asked to refrain from

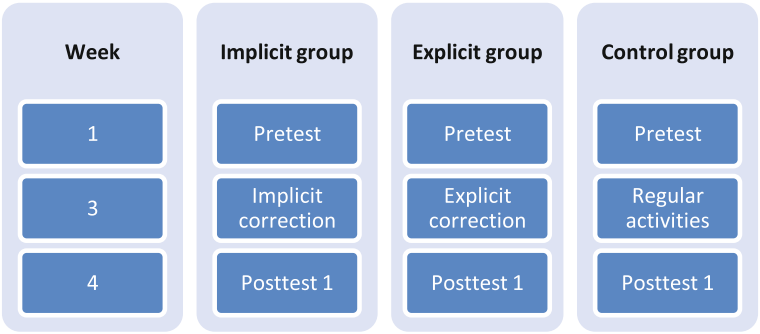


Fig. 1 The design of the study

reacting to errors in the use of the targeted words throughout the experiment. In other words, when these words were mispronounced, no attempt was made to provide corrective feedback or make the students aware of the correct pronunciation in any other way.

The mastery of the pronunciation of the instructional targets was measured on a pretest and a posttest, both of which were administered on the same day in the IG, EG and CG. They included a reading passage containing the targeted words (see Appendix 1) and a task in which the students were asked to summarize the reading passage on the basis of the phrases provided. The first of these, which required controlled production of the targeted items, was envisaged as a measure of explicit knowledge, whereas the second, which called for more spontaneous use of the problematic words, was meant as a measure of largely implicit knowledge. On the reading passage, the students could score 0 or 1 points for each item, with the maximum score of 90 points, and the analysis involved tabulating the means and accuracy percentages for each group on each test. On the communicative task, only the accuracy percentage was computed by dividing the number of accurately pronounced target words by the total number of the instructional targets employed. In both cases, the pronunciation was regarded as correct when it was unlikely to cause a misunderstanding, with the effect that many of the problems involving, for example, imprecise vowel quality or a wrong place of articulation were ignored. Interrater reliability was calculated for a data sample on the pretest and posttest, and the level of agreement reached was satisfactory. The statistical significance of the between-group and within-group differences was established by means of independent samples and paired *t* tests.

5 Research Findings

As can be seen from Table 1, which presents the means, accuracy percentages and the levels of statistical significance, the pedagogic intervention in the form of error correction, irrespective of which corrective technique was used, turned out to be

Table 1 The means and accuracy percentages on the pretest and the posttest

Group	Pretest		Post-test		Differences	
	Reading M/%	Speaking %	Reading M/%	Speaking %	Reading Gain/Sig.	Speaking Gain/Sig.
IG	M = 56.76 63.06 %	62.69 %	M = 68.78 76.42 %	77.03 %	12.02 Sig. 13.36 Sig.	14.34 Sig.
EG	M = 59.06 65.62 %	66.84 %	M = 75.45 83.83 %	78.56 %	16.39 Sig. 18.21 Sig.	11.72 Sig.
CG	M = 59.25 65.83 %	65.87 %	M = 62.15 69.06 %	67.23 %	2.9 N. s. 3.23 N. s.	1.36 N. s.
Sig.	N. s.	N. s.	IG, EG/CG Sig. EG/CG Sig.	IG, EG/CG Sig. EG/CG N. s.		

effective in promoting the development of both explicit and implicit knowledge of the pronunciation of the instructional targets. More precisely, while there were no statistically significant differences ($p < 0.05$) between the experimental groups and the control group on the pretest, such differences could be observed on the posttest, both in the case of the reading task (a difference of 6.63 points between the IG and CG, and 13.3 between EG and CG) and the free production task (a difference of 9.8 between the IG and CG, and 11.33 between the EG and the CG). What is more, the IG and EG improved from the pretest to the posttest on both outcome measures and the gains were statistically significant ($p < 0.05$). In the former they amounted to 12.02 points on the reading task and 13.36 % on the communicative task, whereas in the latter these values were even greater and they stood at 16.39 and 18.21 %, respectively. Even though the control group also did better over time, in all likelihood due to the practice effect, the improvement of 2.9 points and 3.23 % was too small to reach significance.

Moving on to the effectiveness of explicit and implicit feedback, which was the main concern of the present study, although the EG outperformed the IG on both outcome measures, the difference proved to be statistically significant ($p < 0.05$) in reading aloud (6.67 points), but not in free production (1.53 %). This indicates that the use of direct feedback and elicitations coupled with metalinguistic clues was superior to the application of recasts with respect to the explicit knowledge of the pronunciation of the target words, but the value of the CF options in question was comparable in the case of implicit knowledge. It is still noteworthy to point out that the gains in the EG were much greater on both tasks than in the IG, but this is perhaps not surprising given the fact that the students benefitting from explicit error correction were slightly more successful on the pretest than those receiving implicit corrective feedback. While these results stand to some extent in contrast to most of the research findings overviewed in Sect. 4, it should be stressed that explicit correction was still superior when it comes to controlled use of the target words and, as was the case in the research projects conducted by Kim and Mathes (2001), Sanz (2003), and Loewen and Nabei (2007), its effects were not inferior but only comparable to those of implicit CF for spontaneous production. One

possible explanation for such an outcome could be the fact that implicit correction was operationalized not only as recasts, which has been the case in most relevant studies, but also as clarification requests, which are by definition slightly more overt as they require the learner to attempt self-correction, even if the occurrence of the error is never directly mentioned.

There are also other factors that might have affected the findings of the study in somewhat unpredictable ways. In the first place, even though all the instructional targets were included in the role cards for the communicative tasks utilized during the treatment, the frequency of occurrence of the specific items and the number of corrective moves directed at them was not controlled for and, in view of the large number of problematic words, it is clear that some of them must have been used and treated more often than others. Secondly, as the analysis of the recordings and the transcripts of the performance in free production tasks demonstrated, some of the students resorted to avoidance when they were unsure about the pronunciation of specific words, which could have influenced the accuracy percentages, thus affecting the scores of the IG and EG. Thirdly, there are doubts as to the extent to which summarizing a reading passage on the basis of prompts tapped the subjects' implicit knowledge or at least their highly automatized implicit knowledge, since the halting output might have indicated reliance on long-term memory. Moreover, there is evidence for growing awareness of the targeted features, as manifested by a greater number of self-corrections on the posttest. This problem is hard to avoid whenever linguistic knowledge is measured by means of focused communication tasks, although it must be admitted that it can be exacerbated when item- rather than system-learning is the goal. Finally, the effectiveness of the two types of CF could have been the function of the inherent characteristics of specific words, some of which were undoubtedly more difficult than others, as well as individual differences between learners, as it is clear that the working memory capacity, for example, constrains the noticing, interpretation and use of the negative evidence supplied in the corrective move (see Pawlak 2012, for a review). While the impact of the former variable can still be investigated through the reanalysis of the data, the mediating role of the latter could only be examined by designing another study that would take into account selected individual differences factors.

6 Conclusions, Pedagogical Implications and Directions for Future Research

As can be seen from the foregoing discussion, error correction proved to be an effective way of helping learners eliminate persistent errors in the pronunciation of problematic words and it can thus be said to have contributed to item-learning of phonetic features. It is much more difficult to draw definitive conclusions about the contribution of explicit and implicit corrective feedback options, since while the former produced overall greater gains, they conferred a statistically significant advantage on the reading task, aimed to tap explicit knowledge, but not in free

production, intended as a measure of implicit knowledge. Moreover, a number of variables could have affected the findings, such as the frequency of use of specific words and the related intensity of the intervention, the participants' avoidance behavior, the nature of the outcome measure, the variable difficulty of the targeted items as well as individual variation. All of this shows that the results should be considered with circumspection, one should be careful about making far-reaching pedagogical recommendations, and further research is necessary to help us better understand the contribution of different types of corrective feedback and the ways in which this contribution is moderated by intervening variables.

Despite such reservations, it is indisputable that corrective feedback, especially such that is planned, focused, intensive and reflective of learner need, is a viable instructional option in teaching pronunciation features and it should therefore be employed on a regular basis in practical phonetics classes. In view of the fact that it is not entirely clear whether explicit or implicit CF options are more effective and their value may be a function of the nature of the problem and the individual profile of a particular learner, it is perhaps judicious to opt for overt and covert corrective techniques depending on circumstances, or simply combine their use. Another key implication is that system-learning of segmental and suprasegmental phonetic features should be complemented with item-learning of specific words, as this can contribute to the elimination of persistent errors which are frequent in the speech of quite advanced learners despite intensive training in pronunciation they receive. Such an approach appears to be particularly well-suited to practical English classes, where word-level mispronunciations can be responded to with different forms of corrective feedback in the course of communicative tasks, thus enhancing the likelihood of the students gaining greater control of phonetics in free production. Although such guidelines are sound, feasible, and grounded in the available empirical evidence, there is an urgent need for further research into the role of corrective feedback in the development of foreign language pronunciation. Such studies should focus not only on the distinction between explicit and implicit correction, but also on the value of input-providing and output-inducing feedback, not only immediately after the intervention but also in the long run. What is more, an attempt should be made to relate learning outcomes to the inherent characteristics of the targeted features, which may account for the learning challenge and thus constrain the instructional gains, take into account the influence of individual and contextual factors, and probe into the role of learner engagement, as it manifests itself, for example, in different degrees of noticing. Clearly, obtaining insights into the impact of some of these mediating variables may necessitate a combination of quantitative and qualitative methodologies, more reliance on longitudinal research designs, and closer consideration of individual cases.

Appendix 1

The text used as one of the two measures of the mastery of the problematic words. The instructional targets are in bolded italics:

Colonel Frank Robertson was a ***famous*** explorer who had extensive ***knowledge*** of many countries in Africa, ***Asia*** and many ***islands*** in the ***Mediterranean***. He was a partly ***bald*** man with a ***beard*** and a ***moustache***, he ***weighed*** over 100 kg and he had some problems with his right ***knee***. To the surprise of many people, he was ***stingy*** and usually put on old ***clothes***, but often wore a silver ***necklace*** and some other ***valuable jewelry***. He was also very ***calm*** and seldom showed ***signs*** of fear or ***anxiety*** in dangerous situations or gave way to ***basic*** human instincts. He had a ***photographic*** memory which allowed him to memorize ***whole*** books and he was always able to draw ***appropriate*** conclusions. He was remembered for a conference in ***Geneva*** where he ***answered*** questions about ***archeology*** with great ***accuracy*** for ***entire*** 2 days.

Frank Robertson was ***chosen*** the President of the National Geographic ***Society*** three times in a row and his victories over his ***opponents*** and ***competitors*** were always ***decisive***. His accomplishments and ***developments*** were so ***admirable*** at the ***height*** of his career that he was made a ***knight*** by the Queen and ***became*** her advisor. Although Frank was very traditional and hated using the ***computer***, he was extremely ***capable*** and ***successful***. He ***developed*** a number of theories about ***droughts***, ***floods***, ***lightning*** strikes, and other natural disasters that cause ***chaos*** in many parts of the world. He also ***determined*** the location of the ***tomb*** of a legendary Indian prince and the ***scene*** where an ***ancient*** Aztec city used to stand. He led a ***comfortable*** life in California where he woke up at ***dawn*** and ***read*** books ***throughout*** the day, often till ***twilight***. However, every ***July*** and ***August*** or in early ***autumn*** he visited a number of ***foreign*** lands, and ***examined mountainous regions***, ***deserts*** and other far away ***areas*** that might hold important secrets and ***materials*** to be discovered.

He was ***said*** to love ***jeopardy*** and he often told a story about his encounter with a ***jaguar*** as he was ***climbing*** a ***mountain*** in Africa and his ***wound*** from a ***bear paw*** when he was ***exhausted*** after a ***folk*** festival in Alaska. He ***flew*** planes, almost ***drowned*** in the ***ocean*** and jumped with a ***parachute***. At the same time Frank was a very modest person who ***purchased affordable*** things, ***chose*** the cheapest things on the ***menu***, and loved eating ***pears***, ***cauliflowers*** and other ***vegetables***. He was also very ***honest*** and ***righteous***, and often emphasized the ***advantages*** of obeying the ***law***.

References

Allwright, Richard L. and Kathleen M. Bailey. 1991. *Focus on the language classroom: An introduction to classroom research for language teachers.* Cambridge: Cambridge University Press.

Brock, Cynthia, Graham Crookes, Richard R. Day and Michael H. Long. 1986. Differential effects of corrective feedback in native speaker—nonnative speaker conversation. In *Talking to learn: Conversation in second language acquisition*, ed. Richard R. Day, 229–236. Rowley, MA: Newbury House.

Carroll, Susanne and Merrill Swain. 1993. Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition* 15: 357–366.

Chaudron, Craig. 1986. Teachers' priorities in correcting learners' errors in French immersion classes. In *Talking to learn: Conversation in second language acquisition*, ed. Richard R. Day, 64–84. Rowley, MA: Newbury House.

DeKeyser, Robert M. 1998. Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In *Focus on form in classroom second language acquisition*, eds. Catherine J. Doughty and Jessica Williams, 42–63. Cambridge: Cambridge University Press.

DeKeyser, Robert M. 2007. Introduction: Situating the concept of practice. In *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*, ed. Robert M. DeKeyser, 1–18. Cambridge: Cambridge University Press.

Doughty, Catherine J. and Jessica Williams. 1998. Pedagogical choices in focus on form. In *Focus on form in classroom second language acquisition*, eds. Catherine J. Doughty and Jessica Williams, 197–261. Cambridge: Cambridge University Press.

Ellis, Rod. 2001. Introduction: Investigating form-focused instruction. In *Form-focused instruction and second language learning*, ed. Rod. Ellis, 1–46. Oxford: Blackwell.

Ellis, Rod. 2003. *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, Rod. 2006. Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly* 40: 83–107.

Ellis, Rod. 2007. The differential effects of corrective feedback on two grammatical structures. In *Conversational interaction in second language acquisition*, ed. Alison Mackey, 339–360. Oxford: Oxford University Press.

Ellis, Rod. 2009. Corrective feedback and teacher development. *L2 Journal* 1: 3–18.

Ellis, Rod, Sean Loewen and Rosemary Erlam. 2006. Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition* 28: 339–368.

Ellis, Rod. 2010. Epilogue: A framework for investigating oral and written corrective feedback. *Studies in Second Language Acquisition* 32: 335–349.

Jelska-Cydzik, Agnieszka. 2006. Does explicit treatment help students eradicate word stress errors. In *Dydaktyka fonetyki języka obcego*, eds. Włodzimierz Sobkowiak and Ewa Waniek-Klimczak, 93–105. Płock: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej w Płocku.

Kim, He-Rim and Glenn Mathes, G. 2001. Explicit vs. implicit corrective feedback. *The Korea TESOL Journal* 4: 1–15.

Lantolf, James. P. 2006. Sociocultural Theory and L2: State of the art. *Studies in Second Language Acquisition* 28: 67–109.

Lantolf, James P. and Steven L. Thorne. 2007. Sociocultural Theory and second language learning. In *Theories in second language acquisition: An introduction*, eds. Bill VanPatten and Jessica Williams, 201–224. Mahwah, NJ: Lawrence Erlbaum.

Leeman, Jennifer. 2007. Feedback in L2 learning: Responding to errors during practice. In *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*, ed. Robert M. DeKeyser, 11–137. Cambridge: Cambridge University Press.

Li, Shaofeng. 2010. The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning* 60: 309–365.

Loewen, Sean. 2011. Focus on form. In *Handbook of research in second language learning and teaching. Volume II*, ed. Eli Hinkel, 577–592. London and New York: Routledge.

Loewen, Sean and Toshiyo Nabei. 2007. Measuring the effects of oral corrective feedback on L2 knowledge. In *Conversational interaction in second language acquisition*, ed. Alison Mackey, 361–377. Oxford: Oxford University Press.

Long, Michael H. 1991. Focus on form: A design feature in language teaching methodology. In *Foreign language research in cross-cultural perspective*, ed. Kees de Bot, Ralph Ginsberg and Claire Kramsch, 39–52. Amsterdam: John Benjamins.

Long, Michael H. 1996. The role of the linguistic environment in second language acquisition. In *Handbook of research on second language acquisition*, eds. William C. Ritchie and Tej K. Bhatia, 413–468. New York: Academic Press.

Lyster, Roy. 1998. Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition* 20: 51–81.

Lyster, Roy and Hirohide Mori. 2006. Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition* 28: 269–300.
Lyster, Roy and Leila Ranta. 1997. Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition* 19: 37–66.
Lyster, Roy and Kazuya Saito. 2010. Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition* 32: 265–302.
Niżegorodcew, Anna. 2007. *Input for instructed L2 learners: The relevance of relevance*. Clevedon: Multilingual Matters.
Panova, Iliana and Roy Lyster. 2002. Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly* 36: 573–595.
Pawlak, Mirosław. 2004. The role of error correction in teaching pronunciation. In *Dydaktyka fonetyki języka obcego w Polsce*, eds. Włodzimierz Sobkowiak and Ewa Waniek-Klimczak, 66–74. Konin: Wydawnictwo Państwowej Szkoły Zawodowej w Koninie.
Pawlak, Mirosław. 2006. *The place of form-focused instruction in the foreign language classroom*. Poznań–Kalisz: Adam Mickiewicz University Press.
Pawlak, Mirosław. 2007. An overview of focus on form in language teaching. In *Exploring focus on form in language teaching*, ed. Mirosław Pawlak, 5–26. Poznań–Kalisz: Adam Mickiewicz University Press.
Pawlak, Mirosław. 2010. Rola korekty błędów językowych w nauczaniu gramatyki. In *Heteronomie glottodydaktyki: Domeny, pogranicza i specjalizacje nauczania języków obcych*, eds. Zdzisław Wąsik and Aleksandra Wach, 243–260. Poznań: Instytut Filologii Angielskiej.
Pawlak, Mirosław. 2012. *Error correction in the foreign language classroom: Reconsidering the issues.* Poznań–Kalisz–Konin: Adam Mickiewicz University and State School of Higher Professional Education in Konin Press.
Pawlak, Mirosław and Magdalena Pospieszyńska. 2003. Does implicit corrective feedback work for pronunciation errors? In *Dydaktyka fonetyki języka obcego*, eds. Włodzimierz Sobkowiak and Ewa Waniek-Klimczak, 125–138. Płock: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej w Płocku.
Ramirez, Arnulfo G. and Nelly P. Stromquist. 1979. ESL methodology and student language learning in bilingual elementary schools. *TESOL Quarterly* 13: 145–158.
Rosa, Elena M. and Ronald P. Leow. 2004. Awareness, different learning conditions, and second language development. *Applied Psycholinguistics* 25: 269–292.
Russell, Virginia and Nina Spada. 2006. The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In *Synthesizing research on language learning and teaching*, eds. John M. Norris and Lourdes Ortega, 133–164. Amsterdam: John Benjamins.
Saito, Kazuya, and Roy Lyster. 2011. Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning* 62.
Sanz, Cristina. 2003. Computer delivered implicit vs. explicit feedback in processing instruction. In *From input to output: A teacher's guide to second language acquisition.* Mahwah, ed. Bill VanPatten, 241–255. NJ. Lawrence Erlbaum.
Schmidt, Richard. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11: 17–46.
Schmidt, Richard. 2001. Attention. In *Cognition and second language instruction*, ed. Peter Robinson, 3–32. Cambridge: Cambridge University Press.
Sheen, Younghee. 2007. The effect of corrective feedback, language aptitude and learner attitudes on the acquisition of English articles. In *Conversational interaction in second language acquisition*, ed. Alison Mackey, 301–322. Oxford: Oxford University Press.
Sheen, Younghee. 2010. Introduction: The role of oral and written corrective feedback in SLA. *Studies in Second Language Acquisition* 32: 169–179.
Sheen, Younghee and Rod Ellis. 2011. Corrective feedback in language teaching. In *Handbook of research in second language learning and teaching. Volume II*, ed. Eli Hinkel, 593–610. London and New York: Routledge.

Spada, Nina and Patsy M. Lightbown. 2009. Interaction research in second/foreign language classrooms. *In Multiple perspectives on interaction: Second language research in honor of Susan M. Gass*, eds. Alison Mackey and Charlene Polio, 157–175. New York and London: Routledge.

Swain, Merrill. 1995. Three functions of output in second language learning. In *Principles and practice in applied linguistics. Studies in honor of H. G. Widdowson*, eds. Guy Cook and Barbara Seidlhofer, 125–144. Oxford: Oxford University Press.

Swain, Merrill. 2005. The output hypothesis: Theory and research. In *Handbook of research in second language learning and teaching*, ed. Eli Hinkel, 471–483. Mahwah, NJ: Lawrence Erlbaum.

Szpyra-Kozłowska, Jolanta, Iwona Chaber, Grzegorz Pietroń and Sławomir Stasiak. 2004. To correct or not to correct? A study in phonetic correction. In *Dydaktyka fonetyki języka obcego w Polsce*., eds. Włodzimierz Sobkowiak and Ewa Waniek-Klimczak, 124–132. Konin: Wydawnictwo Państwowej Szkoły Zawodowej w Koninie.

Szpyra-Kozłowska, Jolanta and Sławomir Stasiak. 2010. From focus on sounds to focus on words in English pronunciation instruction. *Research in Language* 8:163–174.

Varnosfadrani, Azizollah D. and Helen Basturkmen 2009. The effectiveness of implicit and explicit error correction on learners' performance. *System* 37: 82–98.

Williams, Jessica. 2005. Form-focused instruction. In *Handbook of research in second language learning and teaching*, ed. Eli Hinkel, 671–691. Mahwah, NJ: Lawrence Erlbaum.

An Aptitude for Speech: The Importance of Mimicry Ability in Foreign Language Pronunciation

Martin Hinton

Abstract In recent studies into the factors affecting foreign or second language pronunciation, inborn capabilities, such as the musical ear and the ability to mimic, have received relatively little attention compared to that which has been focused on age, length of study or residence, first language and gender. Since Purcell et al. (1980) included mimicry ability in their list of key factors affecting pronunciation, most research has been aimed at downplaying its importance. This paper has been written as part of a project to design a new style of aptitude test; one which would predict not only success in grammar tests, but also potential communicative proficiency. One step in this process is to attempt to establish a test for aptitude for pronunciation. To that end, the paper sets out to examine to what extent the ability to accurately mimic input after minimal exposure is decisive in determining the overall quality of pronunciation of foreign language learners, in part, because this ability is relatively easy to assess and therefore a good candidate for inclusion in a wider test of communicative aptitude. The paper begins with a critical examination of the research which has been done into the factors affecting pronunciation (Piske et al. 2001) and the role of mimicry and the so-called 'musical ear' (Slevc and Miyake 2006) in the pronunciation of language learners. This review will highlight the lack of an authoritative answer on the question of mimicry and the need for further investigation. There will follow a description of two experiments carried out by the author in which the ability of Polish college students of English to mimic was compared with the quality of their general English pronunciation, and a discussion of the results achieved.

M. Hinton (✉)
NKJO Sieradz, Łódź, Poland
e-mail: mdhinton@tlen.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_8,

1 Introduction

This purpose of this paper is to report on studies which have been carried out by the author into the link between mimicry ability and foreign language pronunciation. These studies were designed, however, not as part of a wider investigation into pronunciation but of one into language learning aptitude. It is important to note this distinction from the beginning, as the focus on mimicry as a potential factor in aptitude testing obviously affects the way in which the experimental work was conducted, and understanding of the true motivation behind the work is essential to properly assess its value. The introductory review of literature, however, will deal largely with issues related to mimicry and foreign language accent and aptitude research will be mentioned only briefly.

2 Background

Aptitude testing has been around since the pioneering work of John Caroll in the 1950s and its ability to predict levels of success on formal language courses has never been in serious doubt. Its usefulness in acquisition environments is a more controversial issue, although a good deal of evidence exists to suggest that those with high aptitude scores do better in all circumstances (Robinson 2005). However, neither Carroll's tests, today known as the Modern Languages Aptitude Test (MLAT) (Carrol and Sapon 1957), nor any of the later innovations on the same theme, such as Pimsleur's Language Battery (Pimsleur 1966), include elements which might assess the ability of candidates to produce speech. The testing is concerned mainly with memory, sensitivity to grammatical distinctions and phonetic coding, that is, the ability to recognise sounds and form associations between them and their written form so that they can later be recalled. There is no aptitude test in common use which requires candidates to produce any sounds: all answers are written on paper, leading to accusations that the tests only predict learner success because of their similarity to learner activities in the classroom: they may only be testing the ability to learn for tests just like themselves.

As the focus of teachers and researchers has shifted away from the ability to reproduce grammatical forms and onto real world communication, so interest has grown in a new concept of aptitude which focuses on the ability to produce comprehensible speech in the target language. Given that degree of foreign accent is one of the key factors distinguishing the oral production of second language learners it seems reasonable to look for innate capabilities which might govern that accent.

Piske et al. (2001) list the following as factors claimed to affect degree of foreign accent: age of learning of L2, length of residence, gender, formal instruction, motivation, language learning aptitude and language use. This last is explained as a variable taking into account time 'spent speaking English with native speakers at home, at work or at school' (Piske et al. 2001, 204)

Under the heading of 'language learning aptitude' they consider two possibilities: musical ability and mimicry. Although they suggest that insufficient study has been done on these factors they make the following observation:

> In summary, musical ability has as yet not been identified as one of those variables that have an important influence on degree of L2 foreign accent. The ability to mimic unfamiliar speech sounds, on the other hand, has repeatedly been identified as a significant and independent predictor of foreign L2 accent (Piske et al. 2001, 202).

The principal work cited in favour of the influence of mimicry is that of Purcell and Suter (1980). They put mimicry ability second only to native language in determining degree of foreign accent and noted, to the disappointment of many educators, that these were factors 'which teachers have the least influence on' (Purcell and Suter 1980, 285). If these conclusions were to be taken at face value, it would lead to a virtual end to pronunciation teaching, something which was unlikely to be accepted by professionals engaged in this activity, and a number of studies have since been conducted to show the efficacy of pronunciation teaching on reducing foreign accent in L2. However, the results of these attempts have been mixed. Piske, MacKay and Flege note that 'Many studies examining the influence of formal instruction on degree of L2 foreign accent have not produced encouraging results for language teachers' (Piske et al. 2001, 200) but, in her review, Derwing is clear that: 'If a second language learner has a desire to sound more like a native speaker, pronunciation instruction can be effective' (Derwing 2003, 551).

On the particular question of whether or not the ear can be trained to identify sounds more effectively through the study of phonetics, the situation is also unclear. A study by Pisoni and his colleagues on the ability to perceive sounds reaches the conclusion that 'there can be little question that laboratory training procedures can be used to selectively modify the perception of nonnative phonetic contrasts' (Pisoni et al. 1994, 160). Attempts by Ashby (2007) to design an experiment to illustrate the usefulness of phonetics courses, however, are less than convincing. Her longitudinal study of students on a year-long phonetics course is successful in showing that isolated vowels are easier to recognise than in-word vowels, but her students actually perform worse in the test after 24 weeks than in the test after 12. The author accepts that there is a 'sizeable drop in achievement' and notes a fall in the median score 'further strengthening the impression of deteriorating standards' (Ashby 2007, 1659). This unhappy result is dismissed as the result of 'ill-matched' tests which do not provide 'a true reflection of level of achievement and a clear demonstration of progression' (Ashby 2007, 1660). Interestingly, she does accept, finally, that 'some learners may be more predisposed towards becoming phoneticians than others. This finding hints at an innate capacity which may not be shared by everyone' (Ashby 2007, 1660).

Even if the ability to improve one's accent through a course of study were conclusively shown, it would not, of course, mean that inherent factors are not important and it is unfortunate that no attempt appears to have been made to assess whether those with innate abilities make better progress on such courses than those without.

Sadly, Jilka (2009a) points to a lack of studies which take mimicry ability to be a 'universal pronunciation skill' (Jilka 2009a, 7) and notes that most research which has been done has involved subjects mimicking sounds in the L2 or attempting to imitate certain foreign accents. Even these trials, however, were less concerned with speakers' ability to mimic and more with perceptual differences between them.

In their work cited above Piske, MacKay and Flege also describe an experiment aiming to rate the importance of the various factors they listed as candidates for influence in foreign language accent. Bizarrely, they chose not to include mimicry ability on the grounds that:

> [v]ariables relating to language aptitude such as mimicry or musical ability […] have usually been found to have no or just a small effect on degree of L2 foreign accent' (Piske et al. 2001, 211–2).

This is in direct contradiction to the passage quoted above from the same paper, but a further quotation may throw light on the change of heart: 'all studies except for one (Flege et al. 1995) have identified mimicry ability as a significant predictor of degree of foreign L2 accent' (Piske et al. 2001, 202). It seems that since Flege's earlier study had found no link, he felt it unnecessary to seek for one again despite acknowledging that others had found one.

A more recent study by Slevc and Miyake (2006) also found a link between musical ability and L2 pronunciation ability. They conclude that:

> The popular conjecture that musical ability is associated with L2 proficiency is not a myth. Although it may be restricted to L2 phonology, individuals who are good at analyzing, discriminating and remembering simple musical stimuli are better at accurately perceiving and reproducing L2 sounds (Slevc and Miyake 2006, 8).

This study offers evidence that aptitudinal factors may have been unfairly passed over in previous experiments as the trend in language learning research swung away from concentration on individual differences and towards other factors. These results are also supported by the work of Narod et al. (2009) who conducted a wide range of tests concerning musicality, language ability and language aptitude as part of the project investigating language talent at the Institute for Natural Language Processing at Stuttgart University. They concluded that musicality correlates strongly with 'productive phonetic talent […] as well as the aptitude for grammatical sensitivity' (Narod et al. 2009, 241): an intriguing claim, since it suggests that the effect of enhanced musicality is not limited to phonetics but is also linked to other linguistic abilities.

It does appear, then, that there is reason to believe that inherent talents can influence second language performance in the area of pronunciation. Conclusive studies linking the ability to mimic unknown sounds with the eventual quality of L2 pronunciation, however, are not easy to find.

3 Method

The possibility of a link between performance in foreign language pronunciation and mimicry ability was investigated in two separate studies using the same methodology. Subjects were given a mimicry test in which they were asked to repeat sounds they had heard only once. Their results in this test were then compared with the results of an assessment of the quality of their pronunciation in English.

The method used in the mimicry test—direct repetition of sounds—is to a degree controversial: it has been argued that it elicits behaviour which is not continued later, once the distance from the original stimulus is increased; but other researchers have been happy to use it (Jilka 2009b). For the purposes of this study there was no controversy: the tests set out to find a correlation between direct mimicry ability and pronunciation performance.

4 Participants

The participants in the two studies were all second year English students at the Teacher Training College in Sieradz. They were aged between 20 and 22, had all learnt English through the Polish school system and none of them had spent significant amounts of time in an English speaking environment other than the College. There was no need, therefore, to consider variance as a result of age of learning, type of teaching, language use or length of residence as, while obviously not identical, all of these factors were reasonably uniform across the group. None of the group had previously studied French, the language used in the mimicry test.

The first group, consisting of 10 students, was assessed at the end of their second year of full-time study in 2008 and the second, of 16 students, at the beginning of their second year in 2010. The pronunciation scores for the first group were collected in real-time during their end-of-year exam where stress and examiner error could have been a factor. The second group were recorded in class and carefully compared, hopefully, leading to a more accurate assessment of their pronunciation relative to one another.

5 Assessment

In order to assess the subjects' mimicry ability they were each played a series of ten French words and short phrases and asked to repeat what they had heard immediately after hearing each one only once. Their responses were recorded and later graded on the following scale: 0 points for no attempt, 1 point for a very poor attempt, 2 points for something recognisable and 3 points for a good pronunciation.

This meant that each candidate received an overall score out of thirty. The candidates were then ranked.

There were two obvious objections to this form of assessment: firstly that it is by its nature subjective, and secondly, that to some degree the personality of the candidate may affect his willingness to repeat an unknown word, especially if he is aware of the fact that he may well not have heard it properly. The first problem was potentially exacerbated by the fact that the assessment was carried out by a native speaker of English with a limited experience of French. In practice, however, these problems did not appear to have a serious effect. The subjectivity of the grading could easily have been checked by using more judges but there seemed little need as many candidates scored 0 or 1 and the only room for controversy was between the awarding of 2 or 3 for better attempts; even this, however, caused few problems and very few scores were changed or even deliberated over on a second listening. The question of an English speaker doing the judging is also an interesting one, as it might be argued that a French speaking judge would compare the candidate's response with his own understanding of how that word should be pronounced rather than strictly with the sound played in the test. Therefore, it may have been an advantage that the judge was not a native speaker of the language used.

Personality, specifically willingness to make some kind of attempt when uncertain, is a more important issue. To a degree, variation was mitigated by the fact that all the candidates were full-time language students, used to speaking in a foreign language. They were also familiar with the examiner and aware of the fact that the quality of their performance would have no consequences for them. More importantly, due to the small size of the sample, it was possible for the examiner to note differences in attitude to the test and take them into account in the final consideration of the results.

Assessment of general English pronunciation was conducted by an experienced examiner, a native speaker of English. In both trials the tasks used to elicit speech were a role-play and a discussion, both done in pairs. In the first experiment candidates were given a percentage score for their pronunciation during their final speaking exam. These scores, however, were treated mainly as an aid to the production of a ranking of the candidates and had no reference to any external standards; in the second trial the candidates were simply ranked by direct comparison with each other until a clear 'league table' emerged—a procedure which can be considered more reliable than the first as it involved repeated listening to the recordings in a controlled environment. The results for each candidate, therefore, can be considered reliable, although the sample size, even taking the repetition of the trial into consideration, is far too small for truly significant conclusions to be drawn.

6 Results

When the rankings for pronunciation and mimicry ability were compared after the first trial the correlation between them was so strong as to be difficult to believe. Of the ten candidates, no fewer than 7 held exactly the same position in both rankings:

the remaining 3 moving by no more than 2 places. The correlation, therefore, worked out at a remarkable $r = 0.96$. What is more, close analysis suggests that this was not a statistical freak: the mimicry scores ranged from 12 to 26 points out of 30 and appeared to reflect genuine differences in the talent to imitate. The candidate who scored 26 was not only a very talented mimic, but was considered by the examiner to have the best English accent he had met with in over a decade of working with Poles, and has won a number of prizes in recitation competitions. The two other candidates who scored above 20 points (22 and 21) were both noticeably weaker than the best mimic and noticeably stronger than the other candidates in their pronunciation. At the bottom of the scale, the candidate with only 12 points from 30 for mimicry clearly possessed the worst English accent.

The wide differences between the mimicry scores meant that questions over the marking of that test were less relevant and no candidates exhibited a marked shyness in giving responses.

The correlation found, however, was both encouraging and worrying: it seemed certain that mimicry ability was exerting an influence on foreign language accent but clearly it was too strong to be a basis for credible conclusions. A repetition of the trial was, therefore, considered essential.

The second trial, conducted in the autumn of 2010 faced one major problem: the overall weakness of the students. Naturally, if there is a link between mimicry skill and accent it would be expected to be relevant at all levels of language ability; though the strength of its influence may vary. The poor level of the students, however, presented certain problems in the gathering of data. Fluency and vocabulary use were both significantly worse than in the original trial, which made assessment of pronunciation more difficult as some candidates produced very little fluent speech and others used very few words. They were also less used to the type of task employed than the first group had been and much less confident generally, which led some candidates to struggle with the mimicry test as well. This problem was made worse by the fact that the two best speakers in the year had to be excluded from the trial as one spoke French and the other had lived briefly in England. Despite these difficulties, the fact that the pronunciation ranks were calculated after multiple listenings to recordings of the conversations of the candidates should mean that the results were more accurate than in the first trial.

The second trial was conducted in much the same way as the first, using an identical mimicry test and scoring procedure, but the material for the pronunciation data was gathered from a class test rather than an exam. The correlation between the rankings for the 16 candidates in the second trial was $r = 0.51$, which seems more realistic but is still impressive. If the genders are separated, the 4 male candidates scored a perfect correlation amongst themselves and the 12 females registered $r = 0.65$. At the top, there was no outstanding mimic with the highest score being 22 out of 30, and no outstanding pronouncer. The best mimic was ranked only 7th for his pronunciation: his mimicry performance, however, may have been boosted by his approach, as he was the only candidate to attempt to assume a French accent and was very bold in making his responses. However, the highest ranked pronouncer also did well in the mimicry test, scoring 20 points. Of

the two candidates with the lowest scores (11) for mimicry, one was somewhat shy in responding during the test and the other, while relatively bold, was simply a very poor mimic. Also, his English pronunciation was easily the worst over the two groups and at times comprehension became impossible because of the strength of his accent.

There were a number of other findings from the trial which are worth recording here. Candidates from the first group were also given a ranking for fluency of speech in English. The correlation between that ranking and the one for mimicry ability was $r = 0.92$. Unfortunately, due to the very poor fluency of many of the second group candidates, no ranking for fluency was attempted in the second trial so little more can be said about this statistic other than that it is interesting and deserves further attention.

Another point of interest was that the average score on the mimicry test was almost identical for the two groups: 16.90 and 16.94. This suggests that the test is capable of producing consistent results and confirms that the mimicry ability of the two groups was very similar even though the overall language performance of the second group was much weaker.

7 Conclusions

The claim, therefore, that mimicry ability is a factor in determining degree of foreign language accent is supported both by the very strong overall statistical data, and by the study of the best and worst scoring individuals. Despite the small size of the samples in both these trials it is difficult to ignore the correlations which were found and not to conclude that in learning environments, other things being equal, mimicry ability is a key predictor of pronunciation skill. These are crucial caveats: this study relates to learning not acquisition environments, where other factors such as length of residence may far outweigh mimicry ability and the participants were chosen specifically to keep other factors as equal possible, so it is impossible to tell the relative strength of mimicry compared to other factors. Certainly, the one other factor which could not be equalised, sex, does appear to be important: in both studies the male candidates occupied the last places in the pronunciation ranking.

From the point of view of aptitude testing, the study has produced results which are very encouraging indeed. Aptitude tests seek to find factors which, other things been equal, predict foreign language achievement, and are not invalidated by the fact that other things will always, in fact, be unequal. The mimicry test was easy to administer and score, and would fit into a test for language production aptitude without difficulty. Mimicry, therefore, appears to be a good candidate for inclusion in the new type of performance-related aptitude testing that researchers in the field have been seeking.

8 Further Research

The trials detailed above can only be considered an introduction to research into the role of aptitudinal factors in foreign language pronunciation. The results achieved would, obviously, be bolstered in their significance by a further repetition on a larger scale and a number of other questions lay themselves open to investigation, such as: do good mimics improve faster on pronunciation courses? Is mimicry ability linked to musical ability? What is the relationship between mimicry ability and traditionally understood language learning aptitude, and, indeed, second language proficiency? And, finally, can mimicry ability be improved by training and would it also improve foreign language pronunciation?

References

Ashby, Patricia. (2007) Phonetics ear-training -Design and duration. ICPhS conference proceedings 2007, 1657–1660. www.icphs2007.de/conference/papers/1281/1281.pdf Viewed: 4/12/2011

Carroll, John and Stanley Sapon. (1957) *Modern languages aptitude test*. New York: Psychological Corporation.

Derwing, Tracey. (2003) What do ESL students say about their accents? *The Canadian Modern Language Review, 59,4 (June)* 547–566.

Dogil, Grzegorz and Susanne Maria Reiterer, Eds (2009) Language talent and brain activity. Berlin—New York : De Gruyter.

Flege, James, Murray Munro and Ian Mackay. (1995) Factors affecting strength of perceived foreign accent in a second language. *Journal of the Accoustical Society of America 97(5) Pt 1 May 1995*, 3125–3134.

Jilka, Matthias. (2009a) Talent and proficiency in langauge. In Dogil, Grzegorz and Susanne Maria Reiterer, Eds, (2009) *Language talent and brain activity*. Berlin—New York : De Gruyter.

Jilka, Matthias. (2009b) Assessment of phonetic ability. In Dogil, Grzegorz and Susanne Maria Reiterer Eds (2009) *Language talent and brain activity.* Berlin—New York : De Gruyter.

Narod, Davide & Susanne Maria Reiterer. (2009) Musicality and phonetic language aptitude. In Dogil, Grzegorz & Susanne Maria Reiterer, Eds (2009) Language talent and brain activity. Berlin—New York: De Gruyter.

Pimsleur, Paul. (1966) *Pimsleur language aptitude battery*. New York: Harcourt Brace Jovanovitch.

Piske, Thorsten, Ian MacKay & James Flege. (2001) Factors affecting degree of foregin accent in an L2 : a review. *Journal of Phonetics (2001) 29*, 191–215.

Pisoni, David, Scott Lively and John Logan. (1994) Perceptual learning of non-native speech contrasts: Implications for theories of speech perception.' pp121-166. In Goodman, Judith & Howard Nusbaum Eds (1994) *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Cambridge, Mass: MIT Press.

Purcell, Edward & Richard Suter. (1980) Predictors of pronunciation accuracy: A reexamination. *Language Learning*. 30/2, 271–87.

Robinson, Peter. (2005) Aptitude and second language acquisition. *Annual Review of Applied Linguistics* (2005) 25. (pp 46–73).

Slevc, L. Robert & Akira Miyake. (2006) Individual differences in second language proficiency: Does musical ability matter? *Psychological Science*, 17(8), 675–681.

MultiROMs and Online Applications Used for Teaching Phonetics in a Secondary School: A Critical Review

Elżbieta Szymańska-Czaplak and Alicja Wujec-Kaczmarek

Abstract The aim of the article is to evaluate the quality of the materials used to teach pronunciation in a Polish secondary school. Although the analysis covers all kinds of aids available for teachers, including a coursebook accompanied by a CD, the authors focus mainly on MultiROMs and online resources connected with a given book. The findings allow the authors to conclude which coursebook sets seem to be most effective in shaping students' pronunciation and which exercises are most helpful, taking into consideration the level of students' English. Additionally, the authors' idea was to check whether these books take into consideration the aspects of pronunciation which prove to be especially difficult for Polish students. The analysis of online resources for teaching and learning pronunciation follows the research conducted by Krajka (English language teaching in the internet-assisted environment. Issues in the use of the web as a teaching medium. Maria Curie-Skłodowska University Press, Lublin, 2007), investigating both segmental and suprasegmental aspects of pronunciation. The aim of the enquiry was to find out in what way the Internet resources can compensate for any shortcomings of traditional coursebooks. There are several reasons why the authors—both secondary school teachers/practitioners—turned to free Internet resources. One of the most important though is the fact that nowadays secondary school students and teachers can seldom afford to buy any additional materials accompanying the foreign language coursebook; thus the search for the resources both available and reliable seems to be essential.

E. Szymańska-Czaplak (✉)
Opole University, Opole, Poland
e-mail: ela@uni.opole.pl

A. Wujec-Kaczmarek
Teacher Training College, Opole, Poland
e-mail: alicjawuka@gmail.com

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_9, © Springer-Verlag Berlin Heidelberg 2013

1 Introduction

Developing student's foreign language competence is a complex process which includes, among others, developing vocabulary, grammar and language skills suitable for different social situations. Discussing the reasons for communication breakdowns, native speakers frequently point to faulty pronunciation rather than grammar mistakes; thus, one of the elements which may contribute to successful performance in L2 is training student's pronunciation.

The acquisition of foreign language pronunciation is influenced by many factors; some of them are connected with school environment and others with external stimuli which reach the student via radio, TV, or direct contact with native speakers. The following analysis is concerned with school environment, in which it depends on the teacher and a coursebook what knowledge of foreign language pronunciation system is conveyed to the student. If the coursebook is building this knowledge systematically, one can expect the student to know the system of vowels and consonants, word and sentence stress rules and some basic intonation patterns of the given language at the end of the given stage of the learning process.

The paper presents the result of analysis of 30 English coursebooks at an elementary and intermediate level, recommended by the educational authorities to Polish secondary schools. The choice of the coursebooks was based on their popularity in Polish schools and in the second part of evaluation on the fact whether they had CD-ROMs attached to student's book or workbook (six MultiROMs/CD-ROMs analyzed). The analysis is focused on the way a given coursebook introduces elements of pronunciation and is an attempt to assess whether exercises appear systematically, which pronunciation elements are practiced most often, what types of exercises are included, and whether the CD-ROM is used adequately.

2 The Analysis of Coursebooks and MultiROMs

There are still some recently published books in the Polish educational market which do not have separate phonetic exercises. They develop all language skills, vocabulary and grammar; they systematically practice spelling but totally neglect pronunciation practice and do not introduce transcription. There is also a numerous group of coursebooks which have a certain element of pronunciation practice, in the form of single exercises distributed at random throughout a book; however, the number of coursebooks whose authors have decided to systematically build the English phonetic system already at the elementary level is not impressive.

Pronunciation exercises included in English coursebooks practice both segmental and suprasegmental elements. Both single sounds and contrasts of several sounds are the focus of attention. Word and sentence stress, as well as intonational patterns constitute another area of development. Coursebooks include exercises on

both weak and strong forms, as well as rhythm. The above mentioned elements are mainly included in student's books and are recorded on class CDs for group practice. Sometimes there are also pronunciation exercises on workbook CDs or MultiROMs for individual practice at home.

Single sounds practice differs from coursebook to coursebook. Sometimes sounds are introduced and practiced one at a lesson, at other times they are contrasted in minimal pairs or sets. The most frequently practiced contrast is the one between high front vowels, long and short. Other minimal pairs appear rarely and the exercises throughout the book do not build the complete system of vowels. Interestingly, the choice of minimal pairs seems to be a random one, some of the contrasts appear in one or two coursebooks only.

Vowels and consonants are practiced in a variety of ways. The most frequent instruction, appearing in all the coursebooks under consideration, is *listen and repeat*. There are also exercises in which students choose the right column to write the word in, the decisive factor being its pronunciation. Some exercises focusing on differences between pronunciation and spelling can also be found in the analyzed coursebooks. They not only show that the same letter can be pronounced in a number of ways but also that some letters do not have to be pronounced at all.

Another group of exercises practices suprasegmental elements. Work on word stress constitutes an essential element of pronunciation practice at both elementary and intermediate level. Most often, students concentrate on stress covering such vocabulary sets as numerals, months, countries and nationalities. Practicing sentence stress is relatively rare at this stage and appears only in few coursebooks. Usually students are asked to repeat the sentence and underline words bearing stress or concentrate on linking words together, using both strong and weak forms of vowels in a given sentence.

Intonation practice can be found very rarely in the elementary coursebooks. In the coursebooks analyzed students are usually asked to repeat the dialogues after the model; however, only few books call the activity intonation practice and emphasize various intonational patterns. The presented examples show that a number of coursebooks introduce elements of pronunciation already at the elementary level. However some of them do it in a very poor and unsystematic way. There are coursebooks which reproduce the same pattern in all units, e.g.: sound contrast practice or tongue twisters; others contain interesting varied pronunciation exercises but do not build the whole system of English sounds. Very few coursebooks provide transcription of words included in mini dictionaries at the back of the book. On the other hand, there are coursebooks which, already at elementary level, start building the complete system of English vowels, diphthongs and consonants, present transcription symbols and even introduce elements of pronunciation to tests. Two such coursebooks are described below.

English in Mind (Puchta and Stranks2006) has been written for lower secondary students who start learning English. This course seems to be unique in comparison to other examined coursebooks as far as teaching phonetics is concerned. The pronunciation material is introduced systematically in Student's Book and practiced extensively in prospective units in Workbook. It is already in Unit 3 when

students start learning phonetic transcription. They learn to recognize phonetic symbols and to write down the pronunciation of every new word in their vocabulary notebooks. They also learn to use a dictionary paying special attention to pronunciation of newly learned words. In Unit 7, when students have learned some longer words, they are taught how to mark stress (in three different ways) and what diacritics mean. In Unit 12 they compare spelling with pronunciation of some English words. In addition to three special units devoted to pronunciation, there is also a pronunciation section in every unit. Having completed the course, the students have a unified knowledge of English phonetic system and they can with a great degree of probability identify letter clusters with proper sounds; they can also recognize word and sentence stress; they know English pronunciation terminology: vowel, consonant, diphthong, word stress, sentence stress, phonetic transcription, IPA and schwa. All the information is clearly marked with the *Pronunciation* heading and the yellow colour reserved for phonetic exercises.

Similar position among upper secondary coursebooks is taken by *English File*. All English sounds and word stress are introduced and intensively practiced in the first year of study. There are exercises which introduce or practice a few sounds in every unit. Every English sound is additionally represented with a specific picture, which facilitates remembering its pronunciation. Sections on pronunciation and word stress are present in all tests and are treated with equal importance to grammar and vocabulary. At the back of the Student's Book there is a sound bank which is referred to extensively in exercises throughout the book. The bank also includes a phonetic domino game, usual spelling for given sounds and a list of words used in the coursebook which have unusual pronunciation. Pronunciation of some British geographic names is also provided. The exercises are diversified but always connected with elements learned by the student at the beginning of the course, making it possible for the students to build up their knowledge and develop pronunciation skills.

As far as variety of activities is concerned, most coursebooks have only repeat *after the cassette* exercises; however, some coursebooks offer activities in which students have to decide what sound they heard and to which category it belongs. Sometimes the instruction is reversed: first students decide and then listen and check, which allows them to consciously practice pronunciation. Another type of activity is choosing the *odd one out* or repeating *tongue twisters*. *Word rhyming* is an interesting exercise on finding identical sounds. In coursebooks for children one can find *songs*, *rhymes* and *chants* for practicing rhythm. However, the variety of activities usually present in primary coursebooks vanishes or changes into monotonous exercises in most secondary coursbooks. Only a few make use of MultiROMs to attract student's attention, provide variety and enable individual work on pronunciation.

However, even the fact that a given coursebook is equipped with a MultiROM does not necessarily mean that pronunciation activities are included. Contrary to the authors' expectations, many coursebook writers do not make any use of the possibilities MultiROMs offer as far as teaching pronunciation is concerned. Krajka (2007) describes numerous ways in which Web-based pronunciation

teaching can contribute to developing students' performance in a foreign language, helping to discriminate sounds, presenting sounds in isolation and context, and practicing their production.

Harmer (2002) enumerates several functions of Web-based instruction, out of which two—*integrated*, when pronunciation is blended with grammar and lexis and *remedial*, when a given phonetic issue is revised—seem to be of utmost importance. Integrating pronunciation teaching with the material covered in the classroom instead of using additional materials for teaching phonetics could encourage teachers to shape their students' pronunciation systematically during lessons. On the other, students working at home could treat MultiROM as a resource tool, which enable choosing exercises which address their individual pronunciation problem and having a patient model for hundreds of repetitions.

With the above assets of computer- assisted pronunciation learning in mind, the authors have made an attempt to evaluate 6 MultiROMs attached to secondary school English coursebooks. Our first choice was free *English in Mind CD-ROM* attached to Workbook 2 and *New English File Intermediate MultiROM PACK* as well as *New English File Intermediate Online*, since both coursebooks scored high in our preliminary general evaluation. However, close examination reveals that the former only partly makes use of the possibilities offered by computers (*Listening and Dictation* section) whereas the latter could be regarded as an example of good practice.

The introduction to *New English File Intermediate* (Latham-Koenig and Oxenden 2006) states the authors' philosophy concerning teaching phonetics, which could be briefly summarized as follows: intermediate students need practice in pronouncing sounds and words clearly; they should be aware of the rules and patterns, be able to use phonetic symbols in their dictionary, be aware of the word and sentence stress; clear intelligible pronunciation should be the goal of the students at that level. Since research shows that the correct pronunciation of individual sounds and syllable stress play a key role in effective oral communication, *New English File* has pronunciation section which prepares students for a speaking activity. It has a system of sound pictures which gives students clear example of words to help them identify and produce the sounds.

The MultiROM attached to the coursebook functions as a MultiROM containing additional language practice when used in the computer and as a CD with extra listening material when used in a CD player. The MultiROM includes Sound Bank section which first shows how to pronounce all English sounds with the help of unique sound pictures system called *English Sounds Pronunciation Chart*. Each picture shows the drawing of a word with a phonetic symbol of a sound incorporated into it and after clicking on the picture a learner can hear first the sound and then the pronunciation of the word. Under each picture there is a short list of example words containing the same sound which a user can listen to. In this way vowels, diphthongs and consonants are presented. Moreover, The MultiROM also includes *Dictation* section with seven dictations helping to practice reception skill. Pronunciation practice in *New English File Intermediate* MultiROM is not connected with the content of the coursebook.

The website of the coursebook http://www.oup.com/elt/global/products/englishfile/intermediate/ also contains pronunciation practice section where a user can again learn the pronunciation of sounds and listen to the examples of words containing the sound by clicking on the picture, listening to the word and the sound, and repeating. The learning process can be made more interesting by playing *Stress Monster Game* which requires identifying the stressed syllable in a word displayed on a gigantic octopus and touching it with an arm of a submarine. The game has three levels of difficulty and makes intriguing sounds; the drawback of the game is the fact that the learner can listen to the word neither before nor after completion of the exercise.

New English File Online Workbook, which has been launched recently by the publishers and can be found at nefonline.pl, requires payment before registering to use the e-learning platform. The online workbook contains a variety of exercises, including pronunciation practice. Additionally, while using the online workbook the user has an access to *Oxford Advanced Learner's Dictionary Online* which is an invaluable tool in practicing pronunciation skills. The exercises are fully interactive, students manipulate the words, drag and drop them, tick them, write them, transcribe them etc. The users can check the answers themselves and the teacher receives the report with the last and the best results the student has achieved. The number of the trials allowed as well as the time in which exercises must be done can be set by the teacher. The technical possibilities given to the teacher by the platform make it incredibly attractive for both sides: the learner and the instructor. Moreover, the whole pronunciation practice is closely connected with the material included in the coursebook. There are seven units, each of which contains a pronunciation section with at least three exercises. The last exercise in each unit concentrates on saying the sounds, words and sentences practiced in the previous exercises.

Unit 1 practices pronunciation of short and long high back vowels and silent letters by the following exercises: a student has to choose the odd-one-out, pronunciation being the decisive criterion; six words are written in phonetic symbols and a student has to write them down in spelling; a student is supposed to tick a silent letter in a word; finally, a learner is asked to practice saying the words containing the sounds that have already been practiced in the unit. The next part of the unit includes practice of long vowels and word stress. In the first exercise, by dragging and dropping, a student is to put the words containing long vowels into the correct column and remaining words into the other column. The next exercise practices recognizing the words with long vowels and yet another one makes a user group the words into three sets according to the position of stress on the first, second or third syllable. What follows is practice on word stress in the words that have been used in the unit. The exercises consist of words divided into syllables and a learner is supposed to tick an accented syllable.

Unit 2 contains two exercises practicing saying and writing the numbers and concentrates on sentence stress. The words in sentences are separated by bullet points and by clicking the right word a student marks the words he/she would normally stress in each sentence. The next part covers working on sounds by

matching the word with the correct sound, writing the words, looking at their phonetic transcription and finally saying them.

Unit 3 concentrates on sentence stress. A student chooses the words he/she usually stresses in the sentences by ticking, writes the words looking at phonetic transcription, pronounces them and repeats the whole sentences paying attention to stressed words. The next part focuses on the sounds spelled as -eigh, -aigh, -ign where a student chooses a word with a different sound out of three words and practices saying them. Further, a student practices sentence stress and individual sounds. An interesting exercise here is the one in which a learner matches the underlined word with a sound illustrated by the picture from a *Picture Sound Chart*.

Unit 4 consists of exercises where a user chooses an odd-one-out sound and marks the stressed syllable in a given word. The next part revises word and sentence stress in the words taken from the unit. Further, a student learns how to distinguish voiceless/s/and voiced/z/used in different sentences, matching the sentence with the sound and again keeps saying the sentences containing these sounds.

Unit 5 revises pronouncing letter combinations -ough and -augh by an exercise based on choosing a word with a different sound and writing a word from the list next to its definition and phonetic transcription. The next activity focuses on the ways of pronouncing *th* sound, where a student chooses the correct pronunciation in the words in the sentence and puts (by dragging and dropping) the underlined words from the sentences into the right column.

Unit 6 introduces practice in pronouncing consonants by matching the words with the correct sounds and then by saying the words containing them. The next part consists of sound recognition exercises, both vowels and consonants are revised.

Unit 7 includes miscellaneous exercises practicing word stress and sounds in isolation, pronunciation of 'schwa', also providing the revision of all sounds.

Another example of good practice is *New Cutting Edge Intermediate* (Cunningham and Moor 2007), where developing pronunciation skills is integrated with teaching grammar and lexis. Student's book includes exercises on stress, weak forms and intonation, while the workbook focuses on problem sounds and word stress. A range of activity types are used, including discrimination exercises and dictation, and an equal emphasis is placed on understanding and reproducing. The CD ROM attached to the coursebook gives an invaluable possibility of listening to the model recordings and then recording yourself. Afterwards a student can listen to both the model and his/her imitation, compare and correct himself/herself. Such an activity seems to be very motivating and helpful in pronunciation practice.

The contents of CD ROM modules corresponds directly to the material covered in the coursebook. Two activities incorporated in *Module 1* cover the topic of people around you and classroom language. The first one practices word stress: a student clicks on the recorded word and after listening to it (as many times as he/she wants) chooses the correctly stressed word out of two answers given. In the second exercise a learner listens to the question, then repeats and records it and listens to his/her recording.

Module 2 contains activities on two tenses: Past Simple and Continuous. In the first exercise a learner listens to a past form of the verb and chooses the correct version, out of two options. The next exercise practices the pronunciation of *was* and *were*. A learner clicks to listen to questions and answers in Past Continuous, then records and plays back each question and answer.

In *Module 3* a student can find two activities, one focusing on the contrast of long and short high front vowels (listen and click on the word you hear) and the other one concentrating on talking about recent holiday. A learner listens to somebody describing his/her holidays (Past Simple sentences) and then records his/her utterance and listens to the recording.

Module 4 has two activities which concentrate on questions with *how long* and *when* and describing people. In the first task a user listens to common questions with *how long* and *when*, then records them and plays back; in the second one he/she listens to descriptions of people and chooses the correct stress in the adjectives used, e.g. *hardworking*, *courageous* etc.

Module 5, with two activities, covers vocabulary connected with work and talking about the future. Again a student chooses the correct stress after listening to a word and then listens to short conversations containing phrases expressing future: learner clicks to listen, then records and plays back both parts.

Module 6, again including two activities, focuses on -ed/-ing endings and extreme adjectives. A student is to listen to the word and click the correct stress pattern; afterwards he/she clicks to listen to the dialogues containing revised words, then records and plays them back.

Module 7, with two activities concerning polite requests and making generalizations with adjectives, is based on listening to the dialogues; then, recording the requests follows, where a student has to pay attention to polite intonation. Then the focus is on listening, recording and playing back the generalizations—paying attention to the word stress on the adjectives.

Module 8 revises compound nouns and linking. First a student listens and clicks on the correct stress pattern; then, he/she clicks to listen to the phrases which are used to talk about machines—paying special attention to how the verbs link. Finally, a student records and plays the utterances back.

In *Module 9* a learner makes predictions and revises the word *make*. In the first exercise the task is to listen to these predictions, then record them and play back; in the second one, a user is to listen to each sentence and type in the phrases.

Module 10 includes the tasks practicing two tenses, Past Perfect and Past Simple, and the revision of adverbs for telling stories. The exercises practice listening to the sentences in the tenses mentioned above and typing them into gaps, with the special attention to the use of short forms. And again student listens to the sentences in past tenses and then records them and plays them back.

Module 11 covers obligation and permission and contains all pronunciation symbols. First, a student listens and chooses the phrase he/she hears and then clicks to listen to the words which are written in phonetic symbols. Afterwards a student records them and plays them back.

Module 12 focuses on *could have/should have/would have* constructions and imaginary situations in the past with *if* -clauses. The exercises comprise listening to model utterances and recording the repeated phrases.

One more interesting element of the New *Cutting Edge Intermediate* CD ROM is a series of *Real life* sections, appearing in each module. They contain video files with comprehension questions, dialogue scripts and the possibility of recording yourself with longer pieces of dialogues, following the examples. CD ROM also includes a phonemic chart with vowels, diphthongs and consonants to listen to and each sound is 'illustrated' by the recordings of a number of typical words containing these sounds. A student, clicking on the phonetic symbol of the sound, can listen to it and if he/she keeps the mouse on the symbol he/she can listen to words containing a given sound.

The remaining MultiROMs analyzed (*Matura Success Intermediate* (Rees-Parnall 2006), *Matura Solutions Intermediate* (Kelly 2007), *Matura Masters Intermediate* (Edwards and Rosińska 2009) do not contain any pronunciation practice. Emphasizing the principles behind the course, the authors of *Matura Success* enumerate a controlled environment for teaching and learning, students and teacher motivation, an interactive approach to learning, thinking training, memorization techniques, an equal emphasis on skills and grammar, a strong focus on vocabulary input and practice, expending general knowledge, building cultural awareness and a principled teaching and evaluation system plus exam preparation and do not even mention the importance of developing pronunciation skills. In *Matura Solutions* there are ten units in the student book and only two pronunciation exercises mentioned in the list of content: intonation when negotiating and emphasizing your opinion. *Matura Masters* offers no Website support although a course *EuroPlus* + *Matura Masters* is advertised as an interactive course using newest technology; however, it is not just supplementary material attached to the coursebook for free but a separate course based on Matura Masters. The above findings concerning Matura exam coursebooks seem to confirm the results of Szpyra-Kozłowska et al. (2003) analysis of phonetic component in examination coursebooks (FC, CAE, CPE), with the conclusion that the exposure to explicit pronunciation exercises in exam coursebooks is none or at least not enough.

3 Conclusions

While choosing a coursebook teachers take various aspects into consideration: they check how the coursebook develops the skills, what grammatical material is included and what vocabulary is introduced. Rarely do they check how pronunciation component is realized. Thus, it is the responsibility of the coursebook authors to provide the most comprehensive development of the student's linguistic competence. Pronunciation practice can be skillfully correlated with the development of other skills, which gives a well-constructed coursebook a big advantage over additional materials that can enrich the coursebook exercises but not replace

them. The special focus on pronunciation practice at elementary level is extremely important because initial pronunciation errors, caused by the teacher's negligence or improper exercises, will be very difficult to eradicate in the further stages of learning.

The evaluation of MultiROMs attached to Workbooks reveals that only some series of coursebooks take full advantage of the new medium as far as teaching pronunciation is concerned. The ones worth recommending are: *New English File Intermediate MultiROM PACK* and *New English File Intermediate Online*, as well as *New Cutting Edge Intermediate CD ROM*. Since every year brings improvements in teaching materials and constantly new editions of well-known coursebooks are published, one may only hope that positive changes will concern pronunciation teaching components as well. MultiROMs treated as resource tools can serve multiple purposes, individualizing student's learning processes, remedying individual pronunciation problems as well as those connected with native language transfer. Having unlimited storage capacity, MultiROMs could provide a huge variety of activities for students to choose from or ignore when not needed because a particular problem does not occur for a speaker of a given nationality. The above ideas would contribute to developing student's autonomy by realizing the concept of coursebook set (MultiROM included) being a resource rather than a set of activities realized by teacher and students during the lesson from cover to cover.

References

Cunningham, Sarah and Peter Moor. 2007. New Cutting Edge Intermediate. Harlow: Pearson Education.

Edwards, Lynda and Marta Rosińska. 2009. *Matura Masters Intermediate*. Oxford: Macmillan.

Harmer, Jeremy 2002. *The Practice of English Language Teaching*. Harlow: Pearson Education.

Kelly, Paul 2007. *Matura Solutions Intermediate*. Oxford: Oxford University Press.

Krajka, Jarosław 2007. *English Language Teaching in the Internet-assisted Environment. Issues in the Use of the Web as a Teaching Medium*. Lublin: Maria Curie-Skłodowska University Press.

Latham-Koenig, Clive and Christina Oxenden. 2006. *New English File Intermediate*. Oxford: Oxford University Press.

Puchta, Herbert and Jeff Stranks. 2006. *English in Mind 2*. Cambridge: Cambridge University Press.

Rees-Parnall, Hilary 2006. *Matura Success Intermediate*. Harlow: Pearson Education.

Szpyra-Kozłowska, Jolanta, Lidia Bąk, Iwona Chaber, Justyna Frankiewicz, Wiktor Gonet, Marta Nowacka, Grzegorz Pietroń and Radosław Święcicki2003. Komponent fonetyczny w podręcznikach przygotowujących do egzaminów Cambridge (FCE, CAE, CPE). In *Zeszyt Naukowy Instytutu Neofilologii (2)*, ed. Włodzimierz Sobkowiak and Ewa Waniek-Klimczak. 137-144. Konin: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej w Koninie.

The English Pronunciation Teaching in Europe Survey: Initial Results and Useful Insights for Collaborative Work

Alice Henderson

Abstract The English Pronunciation Teaching in Europe Survey (EPTiES) is an online survey which aims to collect information about English pronunciation teaching practices in European countries, following the work of Canadian (Breitkreutz et al, TESL Can J, 19:51–61, 2002) and Australian (Macdonald, Prospect 17(3):3–18, 2002) studies. Jointly prepared by teachers in ten European countries (Finland, France, Germany, Ireland, Macedonia, The Netherlands, Poland, Spain, Sweden and Switzerland), the survey has 57 questions organised into nine categories. Initial results for 598 respondents are briefly presented in relation to six categories: characteristics of participants (e.g. age, gender, years teaching experience); teachers' self-evaluation of their own pronunciation skills; teachers' views on the pronunciation-related training they received; assessment of pronunciation; exposure outside the classroom (e.g. subtitled TV, live interaction with native and non-native speakers,); varieties and norms inside the classroom (e.g. for receptive and productive work). Some of the advantages and drawbacks of such collaborative research projects are also addressed.

1 Introduction

In the last decade, English pronunciation teaching has been the subject of several surveys which have provided much useful quantitative and qualitative data. These studies have been done not only in English-speaking countries, such as Canada (Breitkreuz et al. 2002) and Australia (Macdonald 2002) but also in Europe. For

A. Henderson (✉)
Université de Savoie, Chambéry, France
e-mail: alice.henderson@univ-savoie.fr

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_10,

example, Walker (1999) surveyed teachers in Spain and included questions about their training related to teaching English pronunciation. Two other studies focused on this type of training amongst teachers in Britain (Bradford and Kenworthy 1991; Burgess and Spencer 2000). The Council of Europe itself has issued numerous reports related to plurilingual and/or intercultural education, regional and/or minority languages, or language as a subject and/or for teaching subjects, etc.[1] However, no study has specifically addressed how English pronunciation is taught or compared such teaching in several European countries.

The English Pronunciation Teaching in Europe Survey (EPTiES) seeks to fill this gap. This on-line survey is the result of a collaborative effort by teachers in ten European countries: Finland, France, Germany, Ireland, Macedonia, The Netherlands, Poland, Spain, Sweden and Switzerland. As a reaction to a lack of knowledge about what goes on in classrooms across Europe, EPTiES formalises common topics of conversation amongst teachers. The survey is based on two underlying assumptions:

(1) On-line surveys enable people to explain what they do, so they can contribute useful information about what is actually going on in a given context;
(2) Teachers want to know—and benefit from knowing—more about what other teachers are doing.

The time required to carry out such collaborative work should not be underestimated. After initial brainstorming at the Accents 2008 conference in Łodz, Poland, the survey took nearly 2 years to develop. However, in the end the long development span worked to our advantage, as it allowed us to discuss what information would be truly useful, and to propose and modify questions to elicit that information. We were then able to assure uniformity and, especially in the final stages, to correct previously unnoticed typing mistakes. Most importantly, we were also able to draw in more developers from other countries.

Cross-border collaborative work does not necessarily require expensive travel or the most technologically advanced communication tools. In this project, one person is responsible for each country: for developing the questions and, in future, for analysing the data and writing up a book chapter. To date, all discussions have taken place by e-mail, except for a brief meeting of five developers at a conference in June 2009. Video-conferencing or flash meetings have not been used, as they can be difficult to organise from both technological and practical points of view. Although asynchronous interaction between ten people via e-mail can be long and at times frustratingly piecemeal, it has the advantage of leaving time for quiet thought and analysis. This "slow time" feature of asynchronous communication can be beneficial, not just when developing on-line surveys, but also for any collaborative generation of robust new knowledge.

[1] See http://www.coe.int/t/dg4/linguistic/langeduc/le_platformintro_EN.asp

Table 1 The nine categories and the number of questions in each category

Category	Number of questions
Participant information	10
Outside the classroom	2
Pronunciation teaching methods	8
Teaching materials	8
Evaluation of pronunciation	7
Teacher training	3
Views/attitudes	8
Teaching context	7
Model/norm	4

It is important to mention that different countries or states have legal frameworks which may influence the feasibility of such work. At present we are not allowed to use the results for Germany, as our study has not received the necessary ministerial approval.

2 Survey

Teachers at all levels are invited to participate. As of March 4, 2011, 598 teachers had answered the 57 questions,[2] including certain questions which reflect the specificity of national contexts, for example, "At what levels do you teach?" and "Please list your teaching qualifications". The nine categories of questions are presented in Table 1:

Certain *yes–no* questions are followed by a request for more information. The questions about teacher training are open questions, whereas others permit several answers to be chosen from a list, e.g. the two questions in the category "Outside the classroom". The survey was developed via LimeSurvey, a free, open-source application that was chosen for both its user-friendly interface and its basic statistics features, which facilitate analysis of the results. The survey is available on-line.[3] Participants were solicited via personal contacts, mailing lists of professional bodies and "promotional" bookmarks distributed at conferences. Educational institutions and administrative structures were also contacted, in an attempt to reach as many teachers as possible. This data collection phase will continue through the 2011 calendar year.

[2] The paper version of the questionnaire is 29 pages long and is available upon request to the author.

[3] http://vercors.univ-savoie.fr/ressourcesenligne/surveylls/index.php?sid=17898&lang=en.

Table 2 Sub-group, participants per country

Category	Number of respondents
Finland	48
France	23
Ireland	12
Spain	7
Sweden	5
Switzerland	16
Total	111

3 Results and Analysis

The present paper presents a brief analysis of the results obtained up until March 4, 2011 and includes only countries for which there were at least five respondents, excluding Germany. These results (Table 2) therefore represent only 18.5 % of the total, but certain trends are already apparent. When a significant proportion of respondents did not answer a question, this is indicated.

3.1 Characteristics of Participants

The majority of the 598 respondents were female (slightly more than 61 %) and 73 % were non-native speakers of English The average age was 57 years, with an average of 16 years teaching experience. A mere 8 % teach in the private sector.

Among the 111 respondents in the sub-group of 6 countries, 85.5 % were female and their average age was 45. Almost 65 % declared themselves to be non-native speakers of English, a slightly lower fraction than for all the 598 participants (73 %). They had an average of 17 years teaching experience and 17 % taught in the private sector.

Further data are presented for five categories of questions:

- Teachers' self-evaluation of their own pronunciation skills;
- Teachers' views on the training they received;
- Evaluation of pronunciation;
- Outside the classroom;
- Inside the classroom.

For the outside and inside the classroom questions, the results reflect the teachers' points of view, as it was the teachers who indicated what they use and what their learners prefer. This is a weakness of the survey.

3.2 Teachers' Self-Evaluations

Respondents were asked to rate their awareness of their students' skills and goals, and the nature of these goals on five-point scales. Twenty-four percent of respondents did not complete these questions.

Of the teachers who did answer these questions, 40 % judged their own pronunciation to be excellent (5/5) and 29 % expressed high confidence (4/5). However, teachers were less confident about their awareness of their students' goals, with only 10 % claiming excellent awareness (5/5), 38 % claiming good awareness (4/5) and 27 % claiming moderate awareness (3/5). As might be expected, teachers showed greater confidence in their awareness of their students' skills, with 17 % claiming excellent awareness, 43 % claiming good awareness (4/5) and 14.5 % claiming moderate awareness (3/5). Only 4 teachers out of 111 indicated that their students aspire to sound 100 % native-like (Table 3). Three of these four teachers were from France and one was from Sweden. None of the teachers from Finland, Ireland, Spain or Switzerland indicated that their students want to sound native or near-native. The final survey may provide more robust data that will allow us to explain this difference, but it may depend more on the type of course or student and less on the nationality.

3.3 Views on Teacher Training

Detailed replies to this question were provided by 83 of the 111 respondents (Table 4), more than half of whom expressed a lack of satisfaction with the training they had received. An astonishing 20.7 % indicated that their training was "extremely poor" (1/5), 10.8 % indicated it was "very poor" (2/5) and 23.4 % said it was "moderately poor" (3/5). Some mentioned a specific university programme, set books or teachers who influenced them. One teacher's remarks were quite humorous, if depressing: "Not much intelligence, just recycling of tradition, rather like medical doctors practising blood letting for 3,000 years and harming millions in the process". Several mentioned learning the IPA symbols or being self-taught.

In one manner or another, these teachers have developed the necessary skills themselves. This is apparent in their responses to the question asking them to rate the difficulty of teaching English pronunciation (Table 5). Based on Likert scale data, only 3.6 % said they find teaching English pronunciation extremely difficult (1/5) whereas 6.3 % find it extremely easy (5/5), 15.3 % find it very easy (4/5) and a significant 43.2 % find it moderately difficult (3/5).

If teachers have taught themselves, it may be the result of their recognising the importance of pronunciation in relation to the other language skills which compete for teaching time. When asked how important they feel pronunciation is in relation to other language skills (Table 6), 18 % consider it to be extremely important,

36 % consider it very important (4/5) and 17.1 % consider it moderately important (3/5). Only 4.5 % consider it to be not very important (2/5), and less than 1 % consider it to be "not important at all" (1/5).

3.4 Evaluation of Pronunciation

Participants were asked whether or not their assessments of learners' pronunciation were linked to an established scale, for example a national or international scale, and 56 % indicated that this was not the case. Only 29 % of respondents (32 individuals) indicated that a scale was explicitly used; one teacher in Switzerland referred to the BULATS and one teacher in Sweden referred to national tests, as did two teachers in Finland. The other 28 respondents referred to the Common European Framework of Reference for Languages.

Several questions deal with the types and frequency of assessment of learners' pronunciation. In terms of types of assessment, only 30 % of respondents stated that they do an initial diagnostic assessment. Only 7 % of respondents in Finland said that they do an initial assessment and an astonishing 49 % did not answer the question.[4] However, assessments need not be done only at the beginning of a course (Table 7). Assessments can be done during the course and 30.6 % of respondents indicated that this is how they function. Another 30.6 % indicated they use a combination of assessments both during and at the end of the course. It is pedagogically reassuring that a mere 5.4 % indicated that pronunciation is evaluated only at the end of the course. The "other" category was chosen by 18 % of respondents. "Other" may correspond to a different combination of assessments, for example one that includes evaluations at the beginning of the course. In other words, certain participants may have included initial diagnostic types of assessment in their answers to the question about assessment during and/or at the end of the course. It may be that the distinction between these two questions was unclear.

The types of tasks used for diagnostic, formative and evaluative assessment are quite varied (Table 8). Participants could choose more than one from the list of possible answers. The overall proportion of participants choosing the "I don't know" option was quite low across the three types of assessment (0, 2.7 and 1.8 % for diagnostic, formative and evaluative tasks, respectively). However, the proportion of participants who chose the "None of the above" category was only 0.9 % for diagnostic assessments but 4.5 % for evaluative assessment; here the survey should have provided space for an open response. Data from participants in Finland is presented separately (Table 9) as they were directed to a separate question which included Finnish terms for the types of assessment.

[4] However, this last figure should be disregarded, as many more teachers replied to the country-specific questions about types of tasks, as will be shown separately in Table 9.

In terms of the types of tasks used in diagnostic assessments (Table 8), 13.5 % of respondents said they use oral performances, such as short skits, presentations or dialogues. Individual oral exams and reading aloud, with or without preparation, were chosen by 11.7 % of respondents. Oral exams in pairs and listening tasks followed by questions were chosen by 7.2 %. Few participants (3 %) said they use written work for diagnostic assessments. Only 2.7 % indicated that they use another form of diagnostic assessment.

For formative assessments, the respondents indicated that they use the variety of possible tasks more frequently than for diagnostics. The highest percentages are attributed to oral performances (37 % of respondents), listening tasks with questions (36 %) and reading aloud (25.2 %). Just under 20 % use oral exams in pairs in formative assessments, a figure that is only slightly lower (18 %) for evaluative assessments. Participants indicated that they use written work more frequently in formative assessments (11.7 %) and evaluative assessments (13.5 %) than in diagnostic assessments (3 %). This may or may not be linked to work with phonetic symbols. As students show more mastery of the symbols, teachers may find it easier to test certain types of knowledge without evaluating speaking ability. This might also partly explain why oral performances are used less frequently in evaluative tasks (30 %), than they are in formative tasks (37 %). In addition, reading aloud is used less frequently in evaluative assessments (10.8 %) compared with formative assessments (25.2 %). Only 26 % of respondents chose the category "listening followed by questions" for formative assessments, compared to 19 % in evaluative ones. The highest percentage for the "other" choice (7.2 %) was found in evaluative assessments. The fact that participants were not given space to explain this choice is a design flaw of the survey.

Respondents in Finland answered separate, country-specific questions on assessment. For some questions the results for Finland (Table 9) were similar to the results for the overall data. For example the proportion of participants choosing the "I don't know" option was quite low for all three types of assessment: diagnostic (1.8 %), formative (0 %) and evaluative (0.9 %). However, perhaps the most striking difference with the results for the five other countries is found in the "None of the above" category, where the proportion of participants who chose it is for diagnostic assessment (19.8 %) is much higher than the 7.2 % for formative and the 8 % for evaluative assessment. The corresponding figures were much lower for the group of other countries: 0.9, 0.9 and 4.5 %, respectively. The 19.8 % who chose "None of the above" in Finland, may suggest a preference for either using tasks not listed in the survey or for avoiding diagnostic assessments. Given that the Finnish questions used terms in Finnish to reduce any ambiguity, more in-depth analysis of a greater number of responses is required, perhaps including comparison with other country-specific factors, before any conclusions can be made regarding Finnish assessment preferences.

Amongst respondents in Finland, the two most popular forms of diagnostic

assessment were reading aloud (15.3 %) and oral performance (10.8 %). Oral exams in pairs and listening tasks with questions were chosen by only 5.4 % of respondents, and written work was chosen by 2.7 % or respondents. The latter figure is close to the 3 % of respondents from the other countries who chose written work. Individual oral exams seem to be extremely rare across all three types of assessment, as this type of exam was chosen by 0.9 % of respondents for diagnostic assessments, and by 1.8 % of respondents for both formative and evaluative assessment. The figures for individual oral exams are much higher for the other countries: 11.7 % for diagnostic, 24.3 % for formative and 26 % for evaluative assessment.

The respondents in Finland showed a preference similar to the overall survey figures, by choosing oral performances as the most popular choice of task for formative assessments (24.3 %) and then indicating that they use them less in evaluative assessments (18 %). A listening task followed by questions was chosen by 13.5 % of respondents in Finland, compared to only 5.4 % who indicated using it in formative assessments and 9 % in evaluative ones. Reading aloud was also chosen by fewer respondents for diagnostic purposes (15.3 %) than for formative (26 %) or evaluative (21.6 %) ones.

Oral exams in pairs for diagnostics were indicated by 5.4 % of respondents in Finland, but 10.8 % of them indicate using such orals for formative and 18 % for evaluative assessments; eighteen percent is exactly the same proportion as in the overall figures for oral exams in pairs for evaluative purposes. It is, however, worth noting that 26 % of participants from the five other countries chose individual oral exams for evaluative assessments, compared to only 13.5 % of participants in Finland. Other striking differences in the results for evaluative assessment tasks include:

- Written work: only 1.8 % of participants in Finland chose written work, compared to 13.5 % of the other participants;
- Listening and questions: only 9 % of participants in Finland chose listening and questions, compared to 19 % of the other participants;
- Oral performances: only 18 % of participants in Finland chose oral performances, compared to 30 % of the other participants.

3.5 Outside the Classroom

It is a common belief that learners somehow "pick up" the language via their environment, much as children pick up their mother tongue. While this may account for some learning in a native English speaking country such as Ireland, it is difficult to imagine that it can explain learning in the other five countries explored here. The apparent facility with English displayed by certain cultures is frequently attributed to widespread exposure to subtitled films and television programs. This is especially so with reference to northern European countries:

"What else", it is asked, "could explain their seeming superiority in learning English?" In light of research into the role of personality and affective factors (Arnold 1999), we suggest that these "outside the classroom" influences should be seen as merely facilitating opportunities for encountering English; it remains for learners to take advantage of them.

With the goal of investigating this question, teachers were asked to indicate whether or not and to what extent their learners encountered certain sources of English outside the classroom. Perhaps the most surprising result concerns private tuition, where 53.1 % of teachers indicated that their students do not have recourse to such support (Table 10). In addition, the highest percentages for the "I don't know" option were private tuition (23.4 %) and whether or not learners watch news channels (22.5 %). If students are using these sources, their teachers are not aware of it. On the other hand, the respondents were fairly confident in indicating that their students are exposed to subtitled films (64.8 %) and television programmes (51.3 %), and that they have opportunities to use their English (61.1 %). This figure was only very slightly skewed by respondents in Ireland, as these 12 respondents provided only 8 of the 68 *yes* answers for this question. The proportion of incomplete answers was 2.7 % for all of these questions.

In terms of frequency of exposure to English outside the classroom, 45 % of respondents indicated that their students are frequently exposed to subtitled television programmes (Table 11). Even though the data in Table 11 are merely teachers' estimations of their learners' behaviours, not factual accounts from the learners themselves, it is interesting that television should be singled out. The data for films are less categorical (*rarely* = 20.7 %; *sometimes* = 32.4 %; *frequently* = 30.6 %), whereas the contribution of radio was estimated to be less important (*never* = 17 %; *rarely* = 29.6 %; *sometimes* = 24.3 %). On-line resources are important sources, with respondents estimating that 36.9 % of their students *sometimes* use these resources and that 29.7 % of students *frequently* use them. As in Table 10, the proportion of incomplete answers was 2.7 % for all of these questions.

3.6 Inside the Classroom

Sixty percent of the learners were between the ages of 9 and 25 years and 63 % have had 2–9 years of instruction. Sixty-one percent had 1–3 hours of English instruction per week, in groups that were not streamed (43 %) or only slightly streamed (9 %). For the streamed groups, almost 69 % of respondents did not describe how the groups were determined. The 35 teachers who did reply to this question indicated a variety of methods, including written or grammar tests, internationally recognized certifications, short interviews, comprehension tests, self-assessments, or by asking a learner's previous teacher.

The most intriguing results come from the questions about the norms or models of English used by teachers and those preferred by learners. There is not enough space in this paper to explore the differences between countries but this will be

investigated at a later stage in the survey. For each question, respondents could choose one or more of the following categories: *British "RP"*, *GAE*, *Can. Eng*, *Irish Eng*, *Scottish Eng*, *Welsh Eng*, *Austr. Eng*, *NZ Eng*, *South Af. Eng*, *Internat'l Eng*, *Another, No Pref.*, *I don't know*.

The top seven choices for the varieties teachers said they use in the classroom and the varieties teachers felt their learners prefer (in italics) were ranked according to how frequently they were cited by the respondents (Table 12). The questions were:

- "For receptive work (listening, reading), which variety(ies) or model(s) of English do you use in your classes?" and "…of English do your learners generally prefer?"
- "For productive work (speaking, writing), which variety(ies) or model(s) of English do you use in your classes?" and "…of English do your learners generally prefer?"

The results confirm the predominance of two varieties, British "RP" and General American English, in both receptive and productive work. The third rank of "A Type of International English" is surprising as IE remains an imprecise notion. It is true that *The Journal of English as an International Language* has existed since 2007[5] and some recent textbooks do include the words in their title (e.g. *International English Coursebook 1*, Lucantoni and Kellas 2008). However, it is doubtful that among teachers the term IE has as stable a description as RP or GAE. In a Council of Europe discussion paper, Seidlhofer tries to clarify the term:

> 'International English' can be read as shorthand for 'English as an international language' (EIL). The longer term is, however, though more unwieldy, more precise because it highlights the international use of English rather than suggesting, wrongly, that there is one clearly distinguishable, unitary variety called 'International English' (Seidlhofer and Barbara 2003, 8).

The choice "no preference" was the fourth most frequent response for three of the four questions and does not seem noteworthy. The fourth place ranking of Irish and Australian varieties (24.5 %) by teachers for receptive work is more difficult to explain. Although the inclusion of Ireland-based respondents may be responsible for some of this, the importance given to a variety from the other side of the world is puzzling. It might be motivated by local factors, such as teachers from that country or close relations between the teaching institution and an Australian education or business partner. It might also be related to a popular movie or a singer and therefore reflect a faddish interest. A Canadian variety also makes a strong showing for receptive work for reasons that are not clear, with 17 % of teachers choosing it and 4.5 % of learners being said to prefer it. The relatively strong showing of a Scottish variety, chosen by 14.5 % of teachers for receptive work, could easily be explained by teachers' origins, geographical proximity or

[5] http://www.eilj.com/introduction.php, providing "free on-line access to all those involved in the research, teaching and learning of English as an International Language."

institutional links. However, it remains for other research to explore such questions and correlated factors. A prime example is the 2007 nationwide study of Finns' uses of and attitudes to English, which was carried out in order to inform language education policies (VARIENG2011).

4 Conclusion

This article presents a handful of initial results from the EPTiES on-line survey of teaching practices related to English pronunciation. Three main findings stand out:

- Teachers are not satisfied with the training they receive for teaching pronunciation. However, because they see pronunciation as an important skill, they independently acquire enough skills to find teaching pronunciation not extremely difficult.
- Teachers feel relatively confident in their assessments of their learners' skills but they know less about learners' goals.
- Two varieties (RP and GAE) remain predominant in both receptive and productive work, although the presence of a type of English as an international language seems strongly recognised.

Unfortunately, the survey only provides the teachers' perspectives on aspects of learning outside and inside the classroom. This lack of direct information from learners is a weakness of the survey. Future surveys will need to reveal more about learners' perspectives. Moreover, at this stage the number of respondents from each country is too low to allow significant conclusions to be drawn.

As well as presenting the preliminary results of the EPTiES survey, this article shows the value of collaboration as an essential yet vastly underused research mode. Collaborative research promotes certain essential features upon which research communities depend, including a spirit of openness and a genuine desire to share, debate and discuss. However, this mode of research seems less common in the humanities than in the so-called "hard" sciences. The EPTiES survey has shown how rich collaborative work can be. Such pooling of expertise and constructive criticism of each other's ideas can help generate solid results and analyses, which in turn strengthen the research process—for the good of the entire field.

Appendix

See Tables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12.

Table 3 Answers to "To what extent do you feel that your students aspire to have native or near native pronunciation of English?"

	Respondents (%)
1 = Do not at all aspire to sound native	4.5
2	11.7
3	31.5
4	25.2
5 = Aspire to sound native	3.6
Not completed	23.4

Table 4 Answers to "In relation to pronunciation, please rate the teacher training you received"

	Respondents (%)
1 = Extremely poor	20.7
2	10.8
3	23.4
4	19.8
5 = Excellent	6.3
Not completed	19

Table 5 Answers to "For you personally, how easy is it to teach English pronunciation?"

	Respondents (%)
1 = Extremely difficult	3.6
2	8.1
3	43.2
4	15.3
5 = Extremely easy	6.3
Not completed	23.4

Table 6 Answers to "For you personally, how important is pronunciation in relation to other language skills?"

	Respondents (%)
1 = Not important at all	0.9
2	4.5
3	17.1
4	36
5 = Extremely important	18
Not completed	23.4

Table 7 Answers to "Do you evaluate your learners' pronunciation at the end of the course and/or during the course?"

	Respondents (%)
Only at the end of the course	5.4
During the course	30.6
A combination of both	30.6
Other	18
Not completed	14.4

Table 8 Answers to "Which types of tasks do you use for assessment?"

	Diagnostic Respondents (%)	Formative Respondents (%)	Evaluative Respondents (%)
Written work	3	11.7	13.5
Oral performances	13.5	37	30
Individual oral exams	11.7	24.3	26
Oral exams in pairs	7.2	19.8	18
Listening and questions	7.2	26	19
Reading aloud (w or w/o prep. time)	11.7	25.2	10.8
Other	2.7	9	7.2
I don't know	0	2.7	1.8
None of the above	0.9	0.9	4.5

Table 9 Answers from Finland to "Which types of tasks do you use for assessment?"

	Diagnostic Respondents (%)	Formative Respondents (%)	Evaluative Respondents (%)
Written work	2.7	1.8	1.8
Oral performances	10.8	24.3	18
Individual oral exams	0.9	10	13.5
Oral exams in pairs	5.4	10.8	18
Listening and questions	5.4	13.5	9
Reading aloud (w or w/o prep. time)	15.3	26	21.6
Other	3.6	4.5	3.6
I don't know	1.8	0	0.9
None of the above	19.8	7.2	8

Table 10 Answers to "Are students exposed to English outside the classroom?"

	Respondents (%)			
	Yes	No	Some	Don't know
TV subtitled	51.3	21.6	21.6	1.8
Films in cinema subtitled	64.8	10.8	18	2.7
Opportunities to practice English	61.1	5.4	28.8	1.8
Private tuition	2.7	53.1	18	23.4
News channels (BBC, CNN, etc.)	15.3	12.6	46.8	22.5

Table 11 Answers to "Estimate how often students are exposed to English outside the classroom?"

	Respondents (%)				
	Never	Rarely	Sometimes	Frequently	Don't know
Subtitled TV	9.9	16.2	19.8	45	5.4
Subtitled films in cinema	6.3	20.7	32.4	30.6	6.3
Radio	17	39.6	24.3	4.5	10.8
Phone, w NSs and NNSs	15.3	33.3	24.3	9	14.4
Live, w NSs and NNSs	5.4	21.6	43.2	18.9	7.2
On-line resources	1.8	16.2	36.9	29.7	11.7

Table 12 Summary of results from variety(ies) and model(s) questions

Rank	Respondents (%)			
	Receptive, T	*Receptive, L*	Productive, T	*Productive, L*
1	RP 64 %	RP 54 %	RP 61.5 %	RP 50.5 %
2	GAE 55 %	GAE 50.5 %	GAE 39 %	GAE 45 %
3	IE 29 %	IE 15.5 %	IE 17 %	IE 20 %
4	Irish, Austr. 24.5 %	No pref. 5.4 %	No pref. 7 %	No pref. 7 %
5	Can. 17 %	Can., Irish 4.5 %	Irish 6.5 %	Irish, Don't know 4.5 %
6	Scot. 14.5 %	Don't know 3.5 %	Austr. 3.5 %	Can., Austr., Another 2 %
7	No pref. 11 %	Scot. 2.7 %	Can. 2.7 %	–

References

Arnold, Jane (ed.). 1999. *Affect in Language Learning*. Cambridge: Cambridge University Press.

Bradford, Barbara and Joanne Kenworthy. 1991. Phonology on teacher training courses, *Speak Out*, *9* pp. 12–14.

Breitkreutz, Judith A., Tracey M. Derwing and Marian J. Rossiter. 2002. Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19: 51–61.

Burgess, John and Sheila Spencer. 2000. Phonology and pronunciation in integrated language teaching and teacher education. *System*, 28: 191–215.

LimeSurvey 1.90 (open source software). Available at http://www.limesurvey.org/.

Lucantoni, Peter and Lydia Kellas. 2008. *International English Coursebook 1*. London: Hodder Education.

Macdonald, Shem 2002. Pronunciation—views and practices of reluctant teachers. *Prospect*, 17. 3: 3–18.

Seidlhofer, Barbara 2003. A Concept of international English and related issues: from 'real English" to 'realistic English?' (discussion paper). Strasbourg: Language Policy Division DG IV, Council of Europe. Available at http://www.coe.int/t/dg4/linguistic/source/seidlhoferen.pdf.

VARIENG (Research Unit for Variation, Contacts and Change in English, Jyväskylä unit). forthcoming, 2011. Findings of the national survey of English in Finland. Available at http://www.helsinki.fi/varieng/journal/volumes/index.html.

Walker, Robin 1999. Proclaimed and perceived wants and needs among Spanish teachers of English, *Speak Out*, 24 pp. 25–32.

Part III
Phonology

English Vowels: The Long and Short of It

Linda Shockey

Abstract This paper explores the use of the terms "long" and "short" to describe two classes of English vowels with different distributions. We look at phonemic vowel length as it once existed in English and observe that the old quantity distinction does not map comfortably into modern "long" and "short". Alternative terms for the two vowel types are critically discussed and perceptual correlates examined Finally, a pilot study on the pedagogic use of these terms is reported. We conclude that while there is a need to describe such oppositions, the optimal vocabulary for doing so has not yet surfaced.

1 Introduction

It is taken for granted by a large percentage of people who study and teach English that its vowels can be described as 'long' and 'short'. This is an examination of why these terms are used and of to what extent they are appropriate. Heffner (1950, 96) states the situation very clearly:

> Phonological oppositions such as English *heat:hit*, *late:let*, *pool:pull*, *coat:caught* or German *bieten:bitten*, *stehlen:stellen*, *Bruch* (Hose):*Bruch* (Brechung), *Schoss:schoss*, *Hütte:Hütte*, *Sohne:Sonne* suggest that perhaps the vowels [ɪ]. [ɛ], [ʊ], [ɔ] represent types distinct from [i], [e], [u], [o] despite the fact that it is customary to classify [ɪ] with [i], [ɛ] with [e], [ʊ] with [u] and [ɔ] with [o]. Sweet (1877, 8–10) regarded [i] as the narrow form of the high front vowel and [ɪ] as the wide form thereof, because he thought that the body of the tongue was bunched for [i] and more flat for [ɪ]. He made the same distinction for

L. Shockey (✉)
Department of English Language and Literature, The University of Reading, Reading, RG66AA, UK
e-mail: l.shockey@reading.ac.uk

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_11, © Springer-Verlag Berlin Heidelberg 2013

> the other vowel types in question here. Later scholars have substituted the terms tense and lax for narrow and wide, and recognized that this tenseness applied not merely to the tongue muscles but to the entire articulating complex. (Sievers 1901, 98, Sect. 254). Still others have called [i], [e], [u], [o] close (fermées) and [ɪ]. [ɛ], [ʊ], [ɔ] open (ouvertes). Still others incline to treat the second group as short forms of the first, which they call long.

It has thus been common knowledge for more than a century that the two series outlined above can be described phonetically in a variety of different ways, though 'short' and 'long' is arguably the term most frequently used for English. (Dipthongs, which are sometimes regarded as long vowels, will not be covered here).

There is a lack of uniformity as to division of English vowels into two sets (whatever name is chosen). In the next few sections, I will refer to the set containing [i], [u] [ɑ] as 'Set 1' and the one including [ɪ], [ɛ], [ʊ] as 'Set 2'.

Most descriptions of Standard Southern British suggest the following categories. Vowels in the same row are often said to form long/short pairs:

Set 1	**Set 2**
i	ɪ
	ɛ
	æ
u	ʊ
ɔ	ɒ
ɑ	
ɜ	ə
	ʌ

There is also disagreement over [ei] and [əʊ], which sometimes are regarded as phonological monophthongs /e/and /o/ and included in Set 1, e.g. Giegerich (1992, Chap. 3) and sometimes identified as phonetic dipthongs and included in neither class. In addition, Giegerich points out that Standard Southern British (SSB) [ɜ] can be described phonologically as a V + r sequence, which is equally true of American English (AmE) [ɝ].

Within the British tradition, the symbols in Set 1 are often followed by a length mark, as i:, u:, etc.

In the American tradition, some authors think vowels fall into two classes, corresponding to Set 2 vs 'other'. Hockett (1958, 31) and Gleason (1961, 30) both suggest a set of simple vowels including [ɪ,ɛ,æ,ʊ,ə] and a set of complex vowels [ij, ej,aj,oj,uw,ow,aw],[1] the latter being rougly equivalent to Set 1 + the diphthongs.

In contrast, Kenyon (1967, 57) simply lists all vowels in terms of quality and does not include quantity as a distinguishing characteristic. He comments (p. 62) "The words *seat* and *sit* differ in quality of vowel, and if they differ in length, it is not noticed, because they are not distinctive."

Many other descriptions of American English categorise vowels as follows:

[1] Gleason uses [y] instead of [j]

Set 1	Set 2
i	ɪ
(eɪ or) e	ɛ
u	ʊ
(oʊ or) o	ɔ
	ʌ,ə
	æ

[ɔ] and [æ] sometimes appear in Set 1, and [ɑ] is likewise attributed to Set 1 by some writers and to Set 2 by others. [ɝ]/[ɚ] is left out entirely or put into Set 1.

The following sections evaluate terms which have been used to describe these two sets.

2 Short-Long

2.1 Historical Approaches

2.1.1 Segmental Inventories

One argument for using the terms 'long' and 'short' comes from the history of the language, and it is no accident that early scholars of the English language such as Jespersen, Sweet and Jones had an education strongly rooted in philology.

Hogg (1992, 83) and Lass (1994, Chap. 3) note that the Old English vowel system, similar to other West Germanic systems, had a long-short opposition in vowels of identical or very similar quality. Lass suggests the following inventory:

Long system			Short system		
i:	y:	u:	i	y	u
e:	ø:	o:	e	ø	o
a:		ɑ:	æ		

According to Hogg (op cit, 85), there is no evidence in Old English of differences in quality between long and short vowels. "This does not mean to say there was none," he adds, "but if there was, as might be suggested by later developments, it seems to be one of those phonetic details which are immune to our techniques of reconstruction." Cruttenden (2001, 73), however puts forward the following inventory, with differences in quality in many cases:

Long system			**Short system**		
i:	y:	u:	ɪ	y	u
e:	ø:	o:	ɛ	ø	ɔ
æ:			æ		

Regardless of the OE qualities, given the following snapshot of Middle English (ME) vocabulary from about 1,400 (Lass 2000, 69), we can see that that Modern English Set 2 are largely descended from the historical short vowels, though some are derived through shortening of long vowels Liberman (1992, 67). 'deaf', 'good' and 'blood' are examples:

Long	i: bite	**Short**	i bit
	e: meet		e set
	ɛ: meat, deaf		a pass, bat
	ɑ: mate		u cut, put
	u: out		o pot, for
	o: boot, good, blood		
	ɔ: boat		

Many others, some called long (such as the vowels in 'meet' and 'boot' above) are descended from the long vowels, but the picture is not complete because I have not included the ME and modern diphthongs. The mapping between the ME long vowels and diphthongs and the Modern English vowels for SSB and AmE is obviously very complex. In addition, Liberman (loc cit) points out that long vowels can develop historically in English from short vowels through compensation and analogy. The word 'pass' from the list above has a long vowel in SSB.

Modern English vowels described as long and short are hence not straightforward reflexes of the vowels so termed in Middle English, though a rule-governed mapping between ME vowels and modern vowels can generally be found. There are decided remnants of the old quantity system, but the actual contrasts have changed considerably, and differences in quality have become more prominent.

2.1.2 Syllable Shapes and the Mora

Further insight may be achieved by looking at the development of English phonological patterns syntagmatically as well as paradigmatically.

Liberman (op cit, 70ff) comments, "In OE, both *spĕcan* 'speak' and *cēpte* 'kept' were permissible phonetic structures; it follows that a thousand years ago vowel length had not yet come to depend on the type of syllable or the number of postvocalic consonants." Hogg (1992, 118) agrees that originally vowel lenth was entirely phonemic and unpredictable in English, but by about the end of the ninth

century a series of changes had begun which were to continue well into the Middle English period and which all had the effect of tending to make vowel length predictable. Lass (in Blake 1992, 70) adds that these changes reduced the number of environments in which vowel length is contrastive and tending to stablise certain syllable types as 'preferred' or 'optimal'.[2] Anderson and Jones (1977, 163ff) say by the beginning of the fifteenth century these preferred syllable types with respect to possible vowel and consonant combinations were well established. Jones (1989, 115) further suggests that in this period, all stressed syllables had approximately the same duration, so that vowels in open syllables were lengthened. But [because coda consonants were counted as contributing to syllable length], vowels in closed syllables were not. This tendency is noted by Chomsky and Halle (1968, 253), though not in the context of preferred syllable shapes.

Modern English reflects this development in that Set 1 can be found in both open and closed stressed syllables (though not normally in syllables closed by [ŋ] (*[piŋ], *[buŋ])). Set 2 can be found only in closed syllables (*[pɛ], *[bʊ]) if stressed There is thus a constraint against stressed syllables the coda of which includes only vowels from Set 2.

Another way of stating this is that modern English does not allow stressed syllables of one mora.[3] Restrictions on syllable structures in other Germanic languages are often thought to be mora-based, reflecting Liberman's claim (p. 69) that Old Germanic was a mora-counting language. In German (Féry 1997), the coda of monomoraic syllables has a schwa or syllabic sonorant. The coda of bimoraic syllables can have a 'tense' vowel or a 'lax' vowel closed by a consonant. Codas with monomoraic 'lax' vowels are not allowed. Trimoraic or superheavy syllables are the same as the bimoraic ones but with an additional mora.

If we regard Set 1 and 'tense' and Set 2 as 'lax', the pattern of English is very similar to that of German. Set 1 vowels are thus bimoraic in English, Set 2 vowels are monomoraic, and English does not allow monomoraic stressed syllables.

This brief overview is to suggest that defining English vowels as 'short' or 'long' *per se* is not the whole story, as it is in languages like Finnish which have a straightforward distinction between vowels of different quantities. Perhaps, in fact, modern English vowels are better defined in the context of syllable structure, though within this context, it is still clear that two classes of nucleus exist..

Lass (1987, 96) argues in favour of the terms 'long' and 'short': "In most dialects… there are no pairs distinguished solely by length: there are quality

[2] Lass presents an excellent summary of these changes in the Cambridge History of the English Language Volume 2, Chap. 2. They are too many and too complicated to rehearse here.

[3] The mora is an abstract unit of quantity, usually defined within the context of a syllable. Syllable-initial consonants never have moraic weight. As syllable nuclei, short vowels are normally thought to count as one mora, long vowels two. Whether coda consonants add to the mora count depends on the language (Broselow et al. 1997, 47). In Japanese, vowels obey the conventions cited, long consonants add an extra mora and final nasals count as one mora. The word 'Nippon' (Japan) thus has four moras, ni+p+po+n. In English and German, syllabic consonants can be seen as moraic. In Malayalam, consonants cannot carry a mora at all.

differences as well. This has led to many widely-used transcription systems for English indicating only quality at the phonological level, since the length differences can generally be predicted on the the basis of quality…. This is phonetically misleading; and what's worse, makes the incorrect claim that the vowels in bit and beet, for example, are phonologically the 'same kind of animal'."

That they are different animals, says Lass, can be seen in their distribution (as discussed in this section) and by the fact that the same division between long vowels + diphthongs as opposed to short vowels can be found in all varieties of English, even though the qualities may be quite different in different parts of the world.

Yet, this section must lead us to the conclusion that whereas English was once a straightforward quantity language where long and short vowels could contrast in nearly any environment, it has now developed into a language which has a much more complex relationship between vowel quality, vowel quantity, and syllable structure. The terms 'long' and 'short' cannot be used with the confidence which would have been possible a millennium ago.

2.2 Experimental Studies of Duration

Here I assume (along with Broselow et al. (1997)) that durations of English vowels can provide a clue to their phonological length.

2.2.1 British English

Wells (1962) measured the durations of 11 SSB vowels as spoken by 25 males in a h_d environment. The long vowels for this purpose are [i,u,ɔ,ɑ,ɜ] and the short ones [ɪ,ɛ,æ,ɒ,ʊ,ʌ].

Averaged results for some pairs are as follows. The number following the duration measurement (in milliseconds) shows the value of Set 2 vowel/matched Set 1 vowel:

i/ɪ	293/139	0.44
ɔ/ɒ	330/178	0.54
u/ʊ	294/142	0.48
ɑ/ʌ	335/148	0.44
Average of all vowels	300/160	0.53

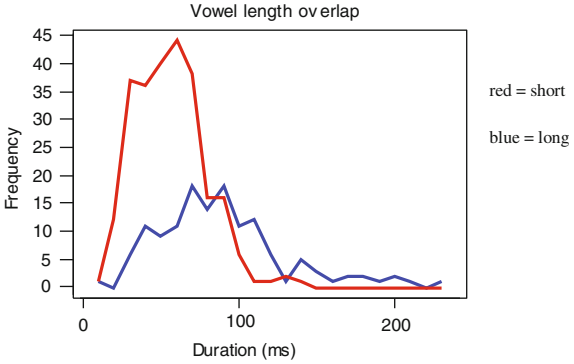
In an attempt to discover whether Wells' ratios can be applied to connected speech, as seen in Umeda's data for American English in Sect. 2.2.1 below, I

extracted the durations of [i] and [ɪ] from a subset of the SCRIBE corpus[4] of British English which included the begin and end times of phonemically-labelled segments as produced in read sentences by ten adult native speakers. The averaged results are not entirely dissimilar to Wells' figures above, allowing for the shorter vowels of connected discourse, though they are, in fact, closer to Umeda's results for American English, below.

	Average duration	**Short-long ratio**	**SD**
i/ɪ	88/55	0.63	Long 40 ms Short 22 ms

The standard deviation is especially striking. This is partly because stressed and unstressed versions of the high front vowels were identified using the same symbol in the database: the [i] in 'very' and 'many' was represented in the same way as the [i] in 'feet'. Correspondingly, the [ɪ] in 'rapid' and 'stupid' was indistinguishable from the vowel of 'tip'.

Even given the unstressed vowels, there is much durational overlap between vowels with the two different labels. It appears, in fact, that the majority of both so-called short and long vowels fall into the 0–100 ms range.



These results mirror those of Lehiste (1970, 37) for Czech vowels in that there is much greater scatter in duration values of long vowels than in short ones. She invokes Trubetzkoy's observation that in a quantity opposition, the short member corresponds to a point in time while the long member is stretchable at will. These data show much more overlap for long/short vowels than Lehiste's Czech data, but hers are from words in isolation, like those of Wells.

Wells (1962, 4) concludes that the length difference is much more obvious in British English than in American English, with very little overlap between short

[4] This is a publically-available database containing samples from four varieties of English. It was a joint venture by CSTR (Edinburgh), University College London, and Cambridge University.

and long vowels. Based on words in isolation, his observation is no doubt correct, but connected discourse appears to present a more complicated picture, at least for i/ɪ. Even so, one suspects that *all else being equal*, [i] would be longer than [ɪ] even in the SCRIBE corpus (cf. Port (1981)), but that it is simply very rare for all else to be equal, due to the large number of linguistic factors (phonetic environment, position in utterance, rate, stress, and more) which determine vowel duration, as observed by Klatt (1976).

2.2.2 American English

House (1961) measured American vowels in a [hə]C_C environment as spoken by three male subjects. Results suggested that the following could be divided into contrastive long-short pairs: [i]-[ɪ], [u]-[ʊ], [e]-[ɛ], and [ɑ]-[ʌ]. He notes that this is sometimes called a tense-lax opposition and puts the other vowels under investigation [æ,ɝ,ɑ,ɔ,o] in the 'tense' class, though noting that the symmetry of the duration system increases if [ɝ] is regarded as lax. There is no overt recognition of the fact that [e] and [o] are normally diphthongs, perhaps due to the belief that all the long/tense vowels of American English have an offglide of sorts, as related in Sect. 1 above.

Peterson and Lehiste (1960) use experimental words in a sentence frame, so the speech represented is connected but controlled. They conclude "It appears rather clear that the vocalic syllable nuclei may be subdivided into classes according to their durations. As a first approximation, the syllable nucleus may be considered as consisting of four short nuclei, [ɪ], [ɛ], [ʊ], and [ə] and nine long syllable nuclei. (The schwa symbol is a more accurate phonetic transcription of the vowel in AmE 'hut' than the [ʌ] normally used for SSB). Set 2 is thus decreased by moving [ɔ] and [æ] into the other category.

Umeda (1975) takes a different approach by measuring the durations of vowels as spoken in a monologue. This means that phonetic environment is not controlled and different numbers of each vowel are found. She concludes that there are four classes of vowels: tense [i, e, o, ɝ], lax [ɪ,ɛ,ɝ], low [æ,ɑ], and diphthongs. Rounded vowels other than [o] are not included in this study. This categorisation seems felicitous in that it bypasses the problems mentioned in Sect. 1 of categorising low vowels into 'long' and 'short' in American English.

The differences found in Set 1 and Set 2 vowels by Umeda (and averaged here from her illustrations) are along the same lines as those of the Peterson and Lehiste study, but are much smaller in magnitude, as can be seen from the following comparative table. The relative difference between the Set 1 and Set 2 vowels is greater in Umeda's results.

	Peterson & Lehiste[a]		Umeda[b]	
i/ɪ	240/180 ms	0.75	93/67 ms	0.72
e(ɪ)/ɛ	270/204	0.76	84/150	0.56
ɑ/ʌ,(ə)	260/230	0.88	125/80	0.64
u/ʊ	260/200	0.77		

[a] Average of five speakers
[b] Average of two speakers in polysyllabic words

Consistent duration differences are thus found between Set 1 and Set 2 vowels in the data examined here.

3 Central/Peripheral

According to Lehiste (1970, 18) vowel duration "appears to be correlated with tongue height: other factors being equal, a high vowel is shorter than a low vowel… It is quite probable that differences in vowel length according to degree of opening are physiologically conditioned and thus constitute a phonetic universal."

It is not always clear which other factors have to be equal: English [i] is shorter than [ɑ] but longer than [ɪ], which is more open, as pointed out by House (1961, 375) perhaps 'peripherality' is called for. A peripheral vowel involves more tongue excursion from a schwa-like position than a more central one and Set 2 vowels are thought to be more central than those of Set 1. Inertia might thus be invoked as an explanation for the differences (it takes longer for the tongue to reach the periphery), which would make durational differences a secondary effect rather than a target feature of the vowel.

F1-F2 charts of American (Peterson and Barney 1952) and British (Deterding 1997) vowels show that [i] and [u] are decidedly more peripheral than [ɪ] and [ʊ]. [ɔ] and [ɒ] make a similar pair for British. It could be argued that [ɑ] and [ʌ] are a peripheral/central pair for AmE, though [æ] and [ʌ] appear to be better candidates in SSB. Vowels such as [ɜ] and [ɝ] do not lend themselves to this type of description, and [eɪ] and [oʊ] also fall outside this scheme, since, like all diphthongs, they have dynamic formants.

It seems highly unlikely that the very large differences reported by Wells above could be entirely attributed to inertia.. It is possible, however, that universal inertial tendencies have been amplified by a language-specific tendency in British English (especially for words in isolation), as is found with the duration of vowels before voiced and voiceless consonants.[5]

Turk et al. (2003) have investigated kinematic characteristics of the tongue blade and jaw during the production of selected vowels by two speakers of American English and find no evidence that that these speakers plan durational

[5] The great majority of the world's languages show a tendency for vowels to be longer before voiced consonants than before voiceless ones in the same syllable: English exaggerates this difference.

differences between tense/lax or higher or lower full vowels. "Our study shows," they conclude "that the acoustic durational patterns commonly observed between American English full vowels can arise as a by-product of kinematic parameters used to encode their quality, and do not necessarily reflect planned durational differences." (p. 27).

Explanations based on inertia thus give some negative answers to whether it is necessary to postulate two distinct sets of vowel targets for English, though a stronger argument can be made for SSB than for AmE.

Chomsky and Halle (1968, 324) use the central/peripheral distinction as part of the definition of tense/lax vowels. Roca and Johnson (1999, 183) also attribute peripherality to tension: "tensing tends to push vowels towards the edge of the vowel chart."

4 Tense/Lax

These terms have supplanted long/short in many online courses on English phonetics, without explanation. Whether they are felt to be more explanatory or simply more up-to-date is not clear. It is generally accepted that experimental corroboration of the feature tense/lax in vowels is difficult, though articulatory features which are potentially quantifiable are suggested by Perkell (see below)

Jones (1957, 39ff) observes for British English, "Those who consider that vowels may be differentiated by degrees of muscular tension distinguish two classes, tense vowels and lax vowels… It is by no means certain that this mode of describing the sounds really corresponds to the facts… The tenseness or laxness of a vowel may be observed mechanically in the case of some vowels by placing the finger on the outside of the throat mid-way between the larynx and the chin. When pronouncing [ɪ], this part of the throat feels loose, but when pronouncing the corresponding tense vowel [i], the throat feels considerably tenser and is somewhat pushed forward… Some teachers get good results by telling pupils to keep their tongues tense or to keep them lax."

Kenyon, speaking of American English, says (op cit, 57) "With some speakers, the tongue is tenser for i,e,u than for ɪ,ɛ,ʊ respectively, and some phoneticians regard this as the important difference between them."

Jones (op cit, 40) notes that it is advisable to apply the terms tense and lax only to close vowels, and this is echoed in Harris (1994, ix), who categorises English high and mid vowels as tense and lax, but excludes low vowels from the dichotomy. Harris later remarks (p. 113) that tense-lax is sometimes referred to as long-short but does not comment on the apparent contradiction that low vowels can be long-short but not tense-lax.

Chomsky and Halle (1968, 324) attribute greater deviation from the rest position, greater distinctiveness, and markedly longer duration to the feature 'tense'. Catford (1977, 206) questions whether such a rest position can be identified unless it is language-specific. Halle and Clements (1983, 7) say, "Tense

vowels are produced with a tongue body or tongue root configuration involving a greater degree of constriction than that found in their lax counterparts; this greater degree of constriction is frequently accompanied by greater length. We note that this feature and (Advanced Tongue Root) are not known to cooccur distinctively in any language and may be variant implementations of a single feature category."

Perkell (1969, 64) says that tense vowels are generally characterised by the attainment of a steady state. In contrast, the lax vowels are characterised by motion or instability. He hypothesises that tense vowels are characterised by greater activity of the extrinsic musculature (the system which moves the tongue as a whole) than for lax vowels. In addition, tenseness could also affect the rigidity of the tongue body (i.e. the intrinsic musculature, inside the tongue) and thereby cause its response to the pull of muscles at its periphery to be less sluggish for tense than for lax vowels.

(The articulatory program for performing) a lax vowel would cause the positioning musculature to contract to a less complete extent with a more flexible tongue body and cause a more transitory and less precise movement of the bulk of the tongue. Absence of the tenseness or rigidity in the tongue body musculature would free the intrinsic musculature to proceed to the following consonant with the minimum delay.

Perkell's conclusions reflect the historical changes noted above in Sect. 2.1.2: "If it is true as Öhman (1963) has suggested (p. 7), that the "(time) unit of natural speech encoding is more the size of a syllable than a phoneme" then a reasonable requirement in the case of English would be that the intrinsic musculature must perform a certain amount of work per stressed syllable. It could accomplish this by contracting to a degree required for a tense vowel or to a degree (which could be the same) required for a lax vowel plus a consonant."

Perkell's theory about the use of the extrinsic musculature in tense vowels is also in agreement with Jones' observations above, since contraction of the mylohyoid and/or the anterior belly of the digastric (extrinsic muscles) may be felt by pressing on the skin under the chin.

Catford (1977, 208), however, remains unconvinced of the utility of [± Tense] for vowels. "For vowels, the existence of such a parameter is dubious, and the use of tense/lax terminology in the phonetic description of vowels is seldom if ever necessary, and should be avoided. This does not mean that the tense/lax terminology is entirely useless in phonological description. Here, tense and lax may, perhaps, be usefully employed as labels to designate phonetically arbitrary classes of vowels that happen to be phonologically distinct. But it should be made quite clear that the selection of terms may be phonetically vacuuous."

The use of tense/lax for English vowels is thus contentious, both because it does not allow for distinctions at all vowel heights and because its phonetic correlates are difficult to discover.

5 Advanced Tongue Root

"ATR", as mentioned above, is thought by some to be equivalent to tense/lax. It is a feature which has been suggested originally to explain vowel harmony in West African languages, many of which have two sets of vowels that are said to differ by one having an advanced tongue root [+ ATR] and the other a retracted tongue root [−ATR]. The [+ ATR] set may also have an enlarged pharynx and a lowered larynx. Words (roots) in these languages must contain vowels that are all of one set or the other. (Fulop et al. 1995).

The two sets of vowels in Maasai (Payne 2003) resemble Set 1 and Set 2 in many respects, although they are not noticeably different in duration:

+ATR	**Neutral**	**−ATR**
i,e,o,u	a	ɪ,ɛ,ɔ,ʊ

This distinction has found a home in English phonology: Harris (op cit, 113) says, "The tense-lax distinction, it is now widely recognised, can be subsumed under…. [ATR] and non-ATR vowels.

Tongue root advancement or retraction as seen in Maasai constitutes a prosody or long component which spreads over a morphological word. Listening to the examples given on the website http:www.uoregon.edu/~dlpayne/maasai/ATR.htm, one can hear that the effect is present throughout the sample words and is especially noticeable in the vowel in the syllable preceding the [ATR] vowel. It may be that such a prosody distinguishes Set 1 and Set 2 in English, but until it is shown that the effect spreads over more than a single segment, it is not an obvious choice as a description for English vowels, even if there is a more concave shape of the base of the tongue for [i] and [u] and a less concave or even convex shape for [ɪ] and [ʊ] (Perkell op cit, 56). The tongue shapes and movements for these vowels are different in other ways, and choosing the tongue root as a distinguisher for English seems a notational convenience in a way that it is not for the African languages.

6 Perception

We have seen that the long/short distinction can be justified from measuring the durations of vowels both in isolation and in connected speech. Here we will investigate whether it is used for recognising English vowels.

6.1 Perception by Natives

6.1.1 Long-Short

Several authors have addressed the question of whether native speakers of English use vowel duration as a distinguishing feature when listening to their own language.

Klatt (1976), in a survey of linguistic influences on segmental duration, mentions an experiment by Nooteboom (1973) in which subjects manipulated the duration of synthetic Dutch vowels until they were thought to sound correct. The durational results echoed the traditional 'long' and 'short' sets in that language. Klatt reports that he performed a similar experiment informally for the words 'bet' and 'bad' in English and got the same result (though see above about whether/æ/is a long vowel in American English). He concludes that it is clear that duration plays an important perceptual role in the identification of English vowels, especially for those pairs which are similar in other respects. This stance is echoed by Nearey (1989).

Other researchers develop the notion that duration may not be a very important cue for native speakers except when vowels are very similar in quality. Bennett (1968) systematically varied duration and quality in synthetic vowels. Results from American listeners "confirmed the hypothesis that the importance of the duration cue is inversely proportional to the distance between the qualities of a given pair of vowels" (p. 65). Ainsworth's (1972) results support this view: "Duration was a relatively more important cue for vowels located in the center of the F1-F2 space" (p. 648).

An area from which results are scarce but provocative is perceptual confusion. An experiment by Miller (1956) showed that perceptual confusions are found amongst members of the long set or the short set, but not across sets. He concludes, (p. 357) "There are two vowel quadrilaterals, the one for the short vowels lying inside the one for the long vowels."

On the other hand, in an early study by Stevens et al. (1969) in which American listeners were asked to identify synthetic steady-state vowels of a fixed duration of 300 ms, they performed to a very high standard. The only vowel which was hard to recognise was [ɪ], and this vowel was not hard for all subjects. Zahorian et al. (1993) produced similar results: "For monophthongal vowels spoken in an isolated-word CVC context, static spectral cues are more important than temporal cues" (p. 1979), though they allow that the situation may be different in continuous speech (p. 1977).

Hillenbrand et al. (2000) tested the importance of American vowel quality vs. duration using 300 /hVd/ syllables. They used four sets of stimuli: naturally-spoken, 'neutral' (adjusted to the average duration of 272 ms), short, and long. They found, like Stevens et al., that the great majority of vowels were identified correctly in all conditions. Some vowel identifications showed a very weak effect of duration (i-ɪ, u-ʊ, ɪ-e-ɛ). The only vowels which were significantly affected were ɑ,ɔ,ʌ and æ,ɛ. Paradoxically, the first three of these fit uncomfortably into the long-short division

(though an ɑ ∼ ʌ pair is suggested in Sect. 3 above) and the latter two are [ɛ], which most agree to call 'short' and [æ], which is ambiguously classified, as seen in Sect. 1, but is often also called 'short'. Another researcher (Bohn 1995, 290) found that vowel duration had little influence on how native speakers differentiated /ɛ/ and /æ/.

Escudero (2001) did a pilot study with speakers of Scottish English and discovered that while they used both quality and duration to signal the difference between [i] and [ɪ] in speech production, in perception the spectral cues were more important than duration: i.e. spectral cues were primary.

6.1.2 Tense-Lax and Others

There is very little in the literature about perception of the tense-lax, peripheral-central, or ATR distinction in vowels. Singh and Woods (1971), when testing for degree of perceptual similarity in vowels, found that 'tense' was a weak predictor of difference. They remark, "Hemdal and Hughes (1967) (who were trying to develop algorithms for computer recognition of vowels) also found difficulty with the feature tenseness, which was… affected by interspeaker variability,"

6.2 Perception (and Production) by Non-Natives

Much evidence exists that non-native speakers of English use duration to recognise English vowels, though perhaps not all in the same way. For example, Morrison (2002) shows that Japanese learners interpret English [i:] as a two-mora version of Japanese [i] and English [ɪ] as the one-mora version. In his words, "Japanese listeners had a duration-based categorical boundary between English /i/ and /ɪ/ that was in the same position as their categorical boundary between Japanese /i:/ and /i/" (p. 112). It can be assumed that speakers of other languages with a quantity distinction which is unrelated to vowel quality will make a similar mapping.

Ingram and Park (1996) found that Japanese listeners use vowel duration to distinguish Australian English [ɛ] from [æ], and Šimáčková (2003) found the same results for Czechs listening to American English. Cebrian's (2002) control group of English subjects relied mostly on spectral cues when distinguishing between tense and lax vowels, whereas Catalan learners resorted mainly to durational differences.

Bohn (1995, 294) suggests that "whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, durational differences will be used to differentiate the non-native vowel contrast."

Bennett (op cit, 78) noted that English speakers use vowel duration as a distinctive feature when judging an unfamiliar contrast, as predicted (or postdicted) by Bohn. Miller and Grosjean (1997, 277) found that French natives used duration when quantity was ambiguous. Escudero and Polka (2003) note that Canadian

English speakers use vowel duration as a way of categorising Canadian French vowels, but observe that this is a counterproductive technique. "… a useful general solution for CE learners of CF is to learn to ignore vowel durations…. (They) must learn to tune into the spectral properties that specify CF vowels." (p. 4)

Fox and Maeda (1999, 4) also argue along these lines, thinking that the tendency of the Japanese to use duration in identifying English vowels, while possibly effective in the short term, stands in the way of native speaker proficiency. "In contrast with the suggestion of Strange et al. (1996) that training include robust duration information, optimal training may in fact not include any duration information, thereby forcing the subject to use more native-like acoustic cues."

They also comment on the resulting effect on speech production (loc cit): "If L2 speakers tend to rely on non-native cues in their categorization of non-native phonemes… this may be related to a similar non-native use of cues in production, i.e. these subjects may distinguish the two vowels /i/ and /ɪ/ based on duration rather than on vowel quality. Therefore… training… which focuses on the proper acoustic cues may be essential not only for subjects to improve their perception but also to achieve more native-like production."

There is some evidence that using duration instead of quality in distinguishing vowels leads to error. Morrison (2002, 8) reports that 73 % of [i]s before voiceless consonants were perceived as [ɪ] by his Japanese subjects. This is because English vowels are significantly shorter before homosyllabic voiceless consonants than before voiced ones, as mentioned in Sect. 3.

It seems clear that the consequences of misarticulating vowel quantity/duration in English are far milder that those of misarticulating quality. The vowel in 'bit' can be stretched out for several seconds without resembling 'beet', and the word 'beet' does not sound like 'bit' even if the vowel is very short. Hess (2002) says that this is not true of standard German. "When I—as a native speaker/listener of German—hear someone pronounce 'Mitte' with a tense [i], it might sound a little overarticulated, but there is nothing wrong with it as long as the vowel is short."

He concludes, "Quality alone is enough to distinguish any two English vowels so that we can establish the whole vowel system in English without needing quantity as a distinctive parameter. Quantity would then be a natural property of some vowels (possibly the tense ones in English). In English [teaching] you can thus afford neglecting quantity in vowels and concentrate on quality alone."

7 Summary

Modern English has two sets of (non-diphthongal) vowels with different distributions. These must therefore be distinguished from each other phonologically, and phoneticians and phonologists have tried to find a feature which will do this. Terms such as long/short, tense/lax, +ATR/−ATR have been suggested. These sets do not have identical members for all varieties of English, though all varieties are thought to have two sets.

While a largely predictable mapping can be found between Old English vowels and modern vowels, a given modern 'long' or 'short' vowel is not always a direct descendent of a vowel of the same quantity. The distinctive use of quantity is part of the West Germanic heritage, but its embodiment in linguistic forms has changed over time and has become inextricably bound up with quality in English. Modern English vowel quantity can be partly attributed to pressure from higher-level linguistic templates and is hence not simply a product of segmental inheritance.

Most of the terms used to distinguish the two sets work very well for the high/close vowels but are more or less problematic for the mid and low vowels, especially in American English. 'Long/short' produces a larger number of reliable pairs for SSB than for AmE.

Duration measurements confirm that all else being equal, the vowels which can we can confidently label as 'short' are consistently shorter than 'long' vowels. But given the very large number of influences on segmental duration, it is quite hard to find a case where members of the two supposed types appear under equal circumstances.

A reliable way to measure tenseness in vowels has not been found, though tests have been proposed. [± Tense] also appears to be most sensibly attributed to high vowels. Peripherality can be determined using F1/F2 plots but, again, cannot be used to distinguish a number of vowels. Advancement of tongue root (ATR) can be observed using various experimental techniques, but this feature may not be appropriate at the segmental level.

As to perception, speakers of English can identify English vowels to a very high standard without knowing their duration. Duration is used distinctively only in marginal cases. However, confusions are either amongst short vowels or amongst long vowels for native speakers, so there is perceptual evidence for the two classes.

It seems there is a natural perceptual technique for all listeners to use vowel duration as a way of distinguishing vowels in cases where quality is difficult to determine. Native English speakers use it to distinguish non-English vowels, and speakers of many other languages use it to determine distinctions in English. Opinions differ as to whether this strategy can be recommended for language learners.

8 Postscript: How is Long/Short Used for Teaching English?

Regardless of whether "long/short" and related terms correctly describe the modern English vowel system, results of an informal survey conducted over the internet in 2012 suggest that these terms are infrequently referred-to in most of today's classes. I had 15 replies from a range of countries. I will not follow the usual referencing style here since I am quoting a collection of e-mails, but some of the names will probably be familiar to you.

Some teachers of English said that their students arrived with the long/short concept already ingrained, presumably a heritage from venerable authors such as

Daniel Jones and A.C. Gimson. P. Carly (UK) says, "the tradition of thinking of certain vowels as forming pairs which differ only in length, as the earlier phonemic transcription seemed to suggest, lives on. Which isn't surprising because your Italians, Spaniards and others can't hear the quality differences, so they assume it must be down to this mysterious factor of 'length' (they are certain that if there was a quality difference, they'd be able to hear it!)"

Most agreed that the terms long/short were not very useful and could, in fact, be misleading. Based on her own survey of the literature, Judy Gilbert (USA) concludes:

1. The terms are not really (phonetically) accurate. It is true that when the sounds are said alone, a long vowel takes longer to say than a short vowel, because the "long" vowels all have a small off-glide /y/ or /w/ sound which takes a little longer to say. Short vowels do not have off-glides and therefore are shorter in duration when said alone. However, sounds are rarely said by themselves, and in the actual context of a word, the duration of the vowel is affected by its surrounding sounds and the stress pattern, all of which can override the intrinsic length of the vowel.
2. The most serious pedagogical objection to using these terms for the distinction between the two vowel categories is that if short and long are used this way, it tends to undermine the meaning of these terms when used to describe the actual duration contrasts which are essential to the English system of stress. Lengthening for stress is important for literacy, as well as intelligibility and listening comprehension, so it would be better to use entirely different terms to describe the vowel contrast in a pair like beat/bit.

Sidney Wood (Sweden) suggests focusing on quality: "My own experience (teaching English in Sweden 40–50 years ago) is that pairs like beat/bit, pool/pull, caught/cot etc. are differentiated by quality (timbre) rather than quantity (length). Swedish does have a quantity distinction, and, apart from some vowel pairs and some regional variation, long and short have the same quality. This is particularly so for Swedish long and short /i/, and if this is carried over into a learner's English it doesn't work for differentiating beat/bit, the result is always e.g. "beat" whether beautifully long or short." A similar point of view is reported by Elcio Souza (Brazil): "The problem here when books teach them as long and short is that Brazilians tend to understand that the sound is simply Brazilian, but pronounced either in a quick way or in a stretched version… That means they get our /i/… and just stretch it or shorten it creating "words" which would be very difficult to identify without the appropriate context."

"Tense/lax" are also suggested and rejected in the exchange, largely because the terms are opaque. Sue Peppe/King-Smith (UK) asks "Does anyone else find the terms 'tense' and 'lax' totally counter-intuitive? Peter Roach (UK) replies, "I often tried asking my students if they had intuitions about "tense" and "lax" as terms for vowel pairs like those 'beat' and 'bid', and never found any agreement about which vowel felt tense and which lax." Petr Rösel (Germany comments) "I avoid the terms tense-lax or fortis-lenis because there is no reliable proprioceptive/kinaesthetic correlation."

Jack Windsor Lewis supports these observations and suggests "sharp" and "soft", based on the usage of Henry Sweet. Other suggestions are "rough/smooth" (Canadian William Acton), "glide/nonglide" (Carole Mawson and Sue Miller (USA)), "out" and "up" (Alexander Selman, UK) and "alphabet" and "relative" (Judy Gilbert, USA).

Helen Fraser (Australia) comments: "It might be worth noting that Australian English is a little different in this respect from the other dialects mentioned so far. It is quite well agreed to be quite well advanced in the process of moving (regressing?) to a more length-based pattern. Despite this, I still don't use length as the main descriptor of vowel phonemes in teaching pronunciation—for reasons similar to those given by Judy Gilbert. In fact I don't like overemphasis on phoneme symbols much at all in pronunciation teaching—especially for teacher education. I think it entrenches the very unhelpful idea that phonemes are 'things' with a single fixed description for each one—where of course each phoneme covers a very diverse range of different phonetic qualities, and the whole set can be categorised in numerous different ways (as we see in this discussion)."

She continues, "I find it far more helpful, in relation to pronunciation, to refer to vowels in terms of the set of words they occur in. John Wells has given us a very helpful reference tool (the HIT vowel, the SHEEP vowel etc.) that allows us to refer to specific vowels without having to commit to an exact (but likely inaccurate) description for it."

This minimalist point of view is supported by Madalena Cruz-Ferreira (Singapore): "I don't think we need terminology at all to teach languages. Language learners want to *use* their new languages, not *talk* about them—unless they're supposed to know linguistic terminology in order to pass easily-marked "proficiency" tests, which is another matter altogether."

9 Pedagogic Conclusion

While there is indubitably philological support for phonemic vowel length in English as a result of its history and membership in the Germanic family, its usefulness for teaching modern English is questionable. My very small sample cannot be said to be definitive (obviously a more systematic investigation is called for), but it seems clear that teachers from many parts of the globe shy away from relying on long/short categories, partly because the concept is not sufficiently transparent and partly because the current trend is away from overt analysis of the system and towards acquisition through use. Nevertheless, for those who use analytical teaching there is a uniform practice of describing the modern English system as having two sets of vowels which share features but are crucially different. The dichotomy is acknowledged, but the vocabulary to describe it has not been standardised.

References

Ainsworth, William A (1972). Duration as a cue in the recognition of synthetic vowels. Journal of the Acoustical Society of America, 51, 648–651.
Anderson, John M. and Jones, Charles, 1977. Phonological Structure and the History of English. North Holland.
Bennett, David, 1968. Spectral form and duration as cues in the recognition of English and German vowels. Language and Speech 11, 65–85.
Blake, Norman (ed)., 1992. Cambridge History of the English Language, Volume 2. CUP.
Bohn, Ocke-Shwen, 1995. Cross-language perception in adults: first-language transfer doesn't tell it all. In Strange, W. (ed.) Speech Perception and Linguistic Experience: Issues in Cross-Language Research, York Press, 273–304.
Broselow, Ellen, Chen, Su-I, and Huffman, Marie, M., 1997. Syllable weight: convergence of phonology and phonetics, Phonology 14, 47–82.
Catford, John Cunnison, 1977. Fundamental Problems in Phonetics. Edinburgh University Press.
Cebrian, Juli, Perception and production of English vowels by native speakers of Catalan: phonetic similarity and acoustic cue reliance. Paper presented at LabPhon 8: Conference on Laboratory Phonology, June 27–30, 2002, New Haven, abstract online at http://www.ling.yale.edu:16080/labphon8/Poster_Abstracts/Cebrian.html.
Chomsky, Noam, and Halle, Morris, 1968. The Sound Pattern of English, Harper & Row.
Cruttenden, Alan, 2001. Gimson's Pronunciation of English. Arnold.
Deterding, David, 1997. The formants of monophthong vowels in Standard Southern British English pronunciation. Journal of the International Phonetic Association 27, 47–55.
Escudero, Paula 2001. The perception of English vowel contrasts: acoustic cue reliance in the development of new contrasts. In Leather, J. and James, A. (eds). Proceedings of the 4th International Symposium on the Acquisition of Second-Language Speech, New Sounds 2000. University of Klagenfurt, 122–131.
Escudero, Paula and Polka, Linda, 2003. A cross-language study of vowel categorisation and vowel acoustics: Canadian English versus Canadian French. Proceedings of the 15th International Congress of Phonetic Sciences, 861–864.
Féry, Caroline, 1997. The mora as a measure of weight and a syllabic constituent, in Bertinetto, P. M., Gaeta, L., Jetchev, D.M. (eds) Certanem Phonologicum III, Papers from the Third Cortona Phonology Meeting, April 1996, Rosenberg and Sellier, 91–110.
Fox, Michelle Minnick, and Maeda, Kazuaki, 1999. Categorization of American English Vowels by Japanese Speakers, Proceedings of the 14th International Congress of Phonetic Sciences, 1437–1440.
Fulop, Sean., Kari, Ethelbert, and Ladefoged, Peter, 1995. Articulatory modelling of so-called advanced tongue root vowels. Abstract 5pSC16 from 129th meeting of the Acoustical Society of America, Washington DC.
Giegerich, Heinz, 1992. English Phonology: An Introduction. CUP.
Gleason, Henry Allen, 1961. An Introduction to Descriptive Linguistics. Holt, Rinehart, Winston.
Halle, Morris and Clements, George N., 1983. A Problem Book in Phonology, MIT Press.
Harris, John, 1994. English sound structure. Oxford: Blackwell.
Heffner, Roe-Merrill S., 1950. General Phonetics. University of Wisconsin Press.
Hemdal, John F. and Hughes, George W., 1967. A feature-based computer recognition program for the modelling of vowel perception. in Wathen-Dunn, W. (ed) Models for the Perception of Speech and Visual Form, MIT Press.
Hess, Wolfgang 2002. Personal communication.
Hillenbrand, James M., Clark, Michael J. and Houde, Robert A., 2000. Some effects of duration on vowel recognition, Journal of the Acoustical Society of America 108, 3018–3022.
Hockett, Charles F. 1958. A Course in Modern Linguistics. Macmillan.
Hogg, Richard (ed), 1992. Cambridge History of the English Language, Volume 1. CUP.

House, Arthur S., 1961. On vowel duration in English. Journal of the Acoustical Society of America 33, 1174–1178.
Ingram, John C.L. and Park, See-Gyoon, 1996. Inter-language Vowel Perception and Production by Korean and Japanese Listeners, Proceedings of The Fourth International Conference on Spoken Language Processing, online at http://www.asel.udel.edu/icslp/cdrom/ vol1/164/a164.pdf.
Jones, Charles, 1989. A History of English Phonology. Longman.
Jones, Daniel, 1957. An Outline of English Phonetics, CUP.
Kenyon, John S., 1967. American Pronunciation. George Wahr.
Klatt, Dennis H., 1976. Linguistic Uses of Segmental Duration in English: Acoustic and Perceptual Evidence" Journal of the Acoustical Society of America 59, 1208–1221.
Lass, Roger 1987. The Shape of English. J.M. Dent & Sons.
Lass, Roger 1994. Old English: A Historical Linguistic Companion. CUP.
Lass, Roger (ed), 2000. Cambridge History of the English Language, Volume 3. CUP.
Lehiste, Ilse 1970. Suprasegmentals. MIT Press.
Liberman, Anatoly, 1992. A birds-eye view of open syllable lengthening in English and the other Germanic languages, North-Western European Language Evolution 20, 67–87.
Miller, George A., 1956. The perception of speech. In Halle, M., Lunt, H., and McClean, H. (eds) For Roman Jakobson. Mouton.
Miller, Joanne and Grosjean, Francois, 1997, Dialect effects in vowel perception: the role of temporal information in French, Language and Speech 40, 277-288.
Morrison, Geoffrey S. 2002. Perception of English/i/and/ɪ/by Japanese listeners. University of British Columbia Working Papers in Linguistics 8.
Nearey Terrance, (1989). Static, dynamic, and relational properties in vowel perception. Journal of the Acoustical Society of America 85, 2088–2113.
Nooteboom, Sieb G., 1973. The perceptual reality of some prosodic durations. Journal of Phonetics 1, 25–45.
Öhman, Sven E.G., 1963. Coarticulation of stops with vowels. Quarterly Progress and Status Report 2/63, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, 1–8.
Payne, Doris L., 2003. Maasai vowels, or what do tongue-root contrasts sound like? Online at http://www.uoregon.edu/~dlpayne/maasai/ATR.htm.
Perkell, Joseph, 1969. Physiology of Speech Production, MIT.
Peterson, Gordon E. and Barney, Harold L., 1952. Control methods used in a study of the vowels. Journal of the Acoustical Society of America 24, 175–184.
Peterson, Gordon E. and Lehiste, Ilse, 1960. Duration of syllable nuclei in English. Journal of the Acoustical Society of America 32. 693–703.
Port, Robert F. Port, 1981. Linguistic timing factors in combination. Journal of the Acoustical Society of America 69, 262–274.
Roca, Iggy and Johnson, Wyn 1999. A Course in Phonology. Blackwell.
Sievers, Eduard, 1901, Grundzüge der Phonetik. Olms.
Šimáčková, Šárka., 2003. Role of duration and quality in nonnative perception and production of English front vowels. http://ifa.amu.edu.pl/plm_old/2003/abs_simackova.htm.
Singh, Sadanand. and Woods, David R., 1971. Perceptual structure of 12 American English vowels. Journal of the Acoustical Society of America 49, 1971, 1861–1866.
Stevens, Kenneth N., Liberman, Alvin M., Studdert-Kennedy, Michael and Ghman, Sven., (1969). Cross-language study of vowel perception, Language and Speech 12, 1–23.
Strange, Winifred., Reiko Akahane-Yamada, R., Fitzgerald, Brett, and Kubo, Reiko., 1996. Perceptual assimilation of American English vowels by Japanese listeners. Proceedings of the International Conference on Spoken Language Processing, Philadelphia.
Sweet. Henry, 1877. A Handbook of Phonetics. Clarenden Press.
Turk, Alice, Matthies, Melanie, Perkell, Joseph, and Svirsky, Mario, unpublished manuscript, 2003. Are American English acoustic vowel duration differences explicitly planned? An articulatory study.

Umeda. Noriko 1975. Vowel duration in American English, Journal of the Acoustical Society of America 58,.434–445.

Wells. John C., 1962. A study of the formants and pure vowels of British English. MA dissertation, University of London. Online at http://www.phon.ucl.ac.uk/home/wells/formants/.

Zahorian, Stephen A. and Jalali Jagharghi, Amir, 1993. Spectral-shape features versus formants as acoustic correlates for vowels, Journal of the Acoustical Society of America 94, 1966–1982.

Quantitative Intervocalic Relations: Evidence for Foot Structure

Tomasz Ciszewski

Abstract The paper discusses intriguing differences between the acoustic characteristics of stressed vowels placed in different prosodic contexts. We have analysed the recordings of speakers of SBrE. The experimental design was as follows. Two male native speakers of Southern British English read 162 mono-, di- and trisyllabic words made of CV sequences. The target items were presented in two contexts: (i) isolated and (ii) phrase-final. The stressed syllables contained all RP vowels and diphthongs, followed by a voiced obstruent, a voiceless obstruent and a sonorant. Vowels were then extracted from target words and their duration, pitch and intensity were measured with PRAAT. Pitch and intensity measurements included three parameters: mean/max values and the intravocalic slope. The total number of observations was 9072. The significance of various types of differences was tested with one-way ANOVA and correlation tests. The results suggest that there exist significant discrepancies between the properties of stressed vowels which depend on the distance between the stressed vowel and the end of the word. These differences on the one hand follow from general laws of speech aerodynamics and on the other serve as important cues in word recognition process.

1 Introduction

Qualitative interdependencies between non-adjacent vowels within lexical items have been observed in a number of languages. These include vowel harmonic processes (van der Hulst et al. 1995), tone assimilation and spreading (Yip 1995) or formant frequencies co-articulation effects (Öhman Sven 1966).

T. Ciszewski (✉)
University of Gdańsk, Gdańsk, Poland
e-mail: angtc@univ.gda.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_12,

In this paper we concentrate on quantitative relations between vowels within initially stressed, morphologically simple words and the ways in which vowel duration interacts with other correlates of stress, namely pitch and intensity. The observed variability of stressed vowel duration indicates that duration is not a reliable correlate of stress in English, since stressed vowels are not necessarily longer than unstressed vowels within the same lexical item. Their durational superiority is often overridden by stress-independent quantitative processes, like pre-fortis clipping (PFC) and final lengthening (FL). Since PFC is contextually conditioned, its effect should be significant irrespective of the prosodic context in which the stressed vowel is placed. Earlier studies (Ciszewski 2010b) have shown, however, that in initially stressed trisyllables produced in isolation stressed vowels in pre-voiced and pre-voiceless contexts do not differ significantly in their durations. As far as the latter regularity (FL) is concerned, its positional conditioning suggests that the durations of word-final vowels should be independent of the number of the preceding syllables in an item and the distance between the stressed vowel and the final unstressed one. The results do not confirm this prediction; word-final vowels in trisyllables prove to be systematically shorter than those in disyllabic words.

The article is organised as follows. In Sect. 2 the experiment design, including the subjects' profile, the stimulus and the measurement criteria, is presented. In the Results and Discussion part we first analyse the differences between the duration of word-final vowels in di- and trisyllables (Sect. 3.1). In the next section durational correlations between stressed and unstressed vowels are discussed which point at the existence of a superordinate durational template whose function is to equalise the total vowel duration within items having a different number of syllables. In Sect. 3.3 the differences in total vowel duration are analysed. Finally, Sect. 3.4 addresses the question of mutual relations between pitch and duration in stressed and unstressed vowels.

2 Experiment Design

Two male speakers of Southern British English took part in a controlled experiment. Each subject read 162 target items (54 monosyllables, 54 disyllables and 54 trisyllables). All items were presented in two contexts: in isolation and phrase-finally (*Say the word…*). Target items were selected according to the following criteria: (i) all monosyllables were of the CVC type, (ii) all di- and trisyllables terminated in [i] (incidentally schwa), (iii) in the stressed vowel position all RP vowels an diphthongs were represented, (iv) the post-stress consonants were of three types: voiceless obstruents, voiced obstruents and sonorants (each vowel and diphthong was placed in all three consonantal contexts), (v) where possible, the initial C was a voiced obstruent. Only vowels were measured in the present study.

Vowel duration was measured with PRAAT (Boersma and Weenink 2005) using waveforms and spectrograms. For vowels followed by consonants, vowel

onset was identified as the point where the target vowel full formant structure was reached and the end of the vowel corresponded to the beginning of the closure phase. The termination of word-final vowels was assumed to coincide with the end of periodic wave accompanied by dispersion of F2/F3.

3 Results and Discussion

3.1 Final Lengthening in di- and Trisyllables

Final lengthening is a regularity whereby the presence of a major syntactic boundary lengthens the word immediately preceding the boundary (for an extensive discussion see: Fletcher 2010). Since in our experiment all items were placed in a pre-boundary position, no significant durational differences should be observed, both as far as whole words and their component vowels are concerned. All polysyllabic target items end in an open syllable, which –except for a few isolated cases– contains the short vowel [i]. In isolated and phrase-final position this vowel is known to undergo a lengthening process known as 'happy tensing' (Fabricius 2002).

The significance of the difference in duration of word-final vowels in di- and trisyllables was tested for both subjects with one-way Anova (alpha 0.05; $n = 108$). The results point to a highly significant effect of the number of syllables in an item on the duration of the word-final vowel (S1 $p = 2.5E - 40$; S2 $p = 2.5E - 44$). Thus, FL proves to be sensitive to the overall duration of the word, or more precisely to the distance to the opposite edge of the word. Therefore, the positional motivation of the processes is overridden by some sort of non-local conditioning which controls the degree to which the process affects its target vowels. The differences in final vowel duration in di- and trisyllabic words are illustrated in the graph below (Fig. 1).

As far as the degree of FL is concerned, the only conceivable explanation for the variation observed between 2- and 3-syllable words seems to be the total duration of preceding (non-final) vowel(s). It has been observed that the accumulated duration of V_1 and V_2 in trisyllables is systematically greater that the duration V_1 in disyllables. Simultaneously, final vowels in trisyllables are shorter than final vowels in disyllables. While the former regularity is obvious (given the fact that mean V_2 duration in trisyllables is greater than the mean difference between V_1 duration in di- and trisyllables), the latter one concerning word-final vowels remains entirely accidental unless it is interlinked with the former. Thus, in order to check whether the duration of word-final vowels is related to the duration of non-final vowels, one needs to compare the differences between (i) V_1 duration in disyllables and the accumulated ($V_1 + V_2$) duration in trisyllables on the one hand with (ii) the differences between final vowel durations in the two groups of words on the other. Our assumption is that the two differences should be

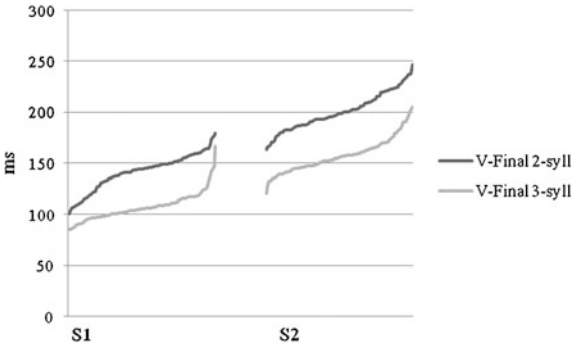
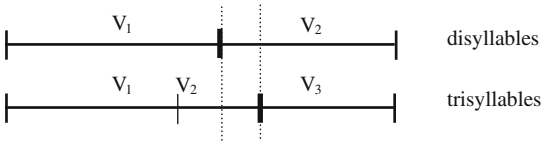


Fig. 1 Duration of final vowels in 2- and 3-syllable items

comparable. This is schematically illustrated below. (The interval corresponding to the differences in question is enclosed between the dashed lines.).



Ideally, the two differences should neutralise each other, i.e. the difference between them should be close to zero. One has to remember, though, that (i) the duration of stressed vowels varies significantly due to the differences in their phonemic length (especially in disyllables) and the influence of PFC and (ii) the total vowel duration in di- and trisyllabic words also displays a certain amount of variation. Thus, it is unlikely that the differences between the durations of non-final vowels and the differences between the durations of final ones in the two groups of items will be identical. However, mean differences between the durations of non-final and final vowels prove negligible.

Mean differences: non-final vs. final V (ms)	S1	S2
A. $(V_1 + V_2)^{TRISYLLABLES}$-$V_1^{DISYLLABLES}$	20.9	22.0
B. $V_2^{DISYLLABLES}$-$V_3^{TRISYLLABLES}$	35.3	43.3
Mean A—Mean B	14.4	21.3

Apart from the difference between the mean values (A-B above), we have also calculated the differences between the durations of non-final vowel(s) in di- and trisyllabic items with a phonemically identical V_1 placed in the same PFC context, e.g. *biddy* ∼ *bigamy*, and compared them with the difference in final vowel duration for each pair of items. Graph 2 below illustrates the results. (Subject-individual differences between $(V_1 + V_2)^{TRISYLLABLES}$-$V_1^{DISYLLABLES}$ and $V_2^{DISYLLABLES}$-$V_3^{TRISYLLABLES}$ have been arranged in ascending order and are labelled

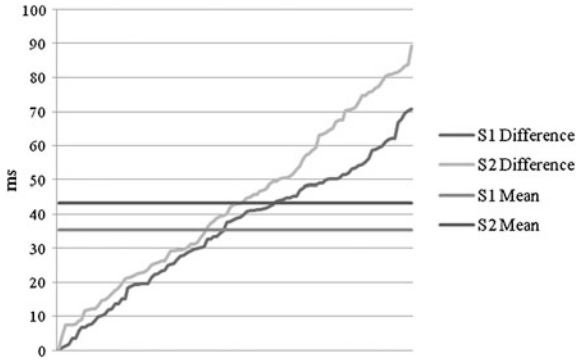


Fig. 2 Differences between the duration of non-final vowels versus the differences between final vowels in di- and trisyllables

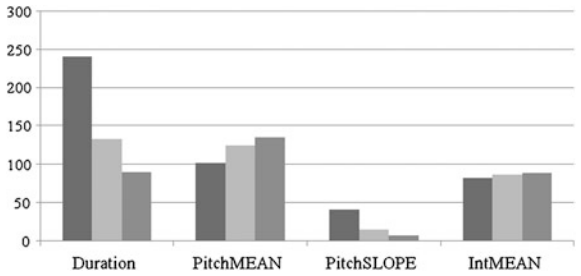


Fig. 3 Differences in stressed vowel duration, mean pitch, pitch slope and mean intensity in monosyllables (*bid*), disyllables (*biddy*) and trisyllables (*bigamy*) (S2)

as 'Difference' for convenience; mean values of these differences (S1 = 35.3 ms and S2 = 43.3 ms) are represented by horizontal lines.)

Surprisingly, despite considerable inter- and intra-speaker variation, the durational differences between V_1 in disyllables and (V_1 + V_2) in trisyllables on the one hand, and those between the final vowels in di- and trisyllables on the other are nearly identical for both subjects, which shows that the relation between them is almost perfectly proportionate, i.e. regardless of speaker-individual differences in the absolute duration of final and non-final vowels, the differences between di- and trisyllabic words are constant and amount roughly to 30 $\sim$ 40 ms.[1] This is a much more realistic result than the proposed 'zero difference' (note that the difference is a derivative of a number of variables that affect V_1 duration (phonemic length, PFC effects), V_2 duration in trisyllables (weak PFC effect) and the natural intra-speaker variation in final vowel duration. The remaining non-reducible 30 $\sim$ 40 ms difference may then be interpreted as being mechanistically conditioned by insurmountable articulatory requirements on V_1 duration which are imposed by the following consonant and, to a some extent by its intrinsic phonemic length.[2] This, in turn, suggests that not only is the degree of FL conditioned

[1] Which corresponds to approximately 10 % of total vowel duration. This result is supported by a number of analyses (for an overview of studies on just-noticable durational differences see: Waniek-Klimczak 2005).

by total vowel duration, but that it is also primarily related to the duration of the stressed vowel. Hence, FL and PFC must also be interdependent.

In conclusion, the differences in degree of FL in 2- and 3-syllable words are indeed coupled with the differences in the accumulated duration of the preceding vowel(s). The 'equilibrium' is not perfect, though. It is disturbed by a complex network of durational interrelations between phonemic length of V_1 and the PFC context (which has an effect on both V_1 in disyllables and V_2 in trisyllables) and partly by the natural variation in the duration of final vowels.

3.2 Durational Correlations Between Stressed and Unstressed Vowels

Bearing in mind the complex combination of factors that influence V_1 duration, the lengthening and tensing of V_2 and the global pre-pausal lengthening of the whole word (which to some extent also affects V_1), the possibility of a systematic relation between V_1 and V_2 in disyllables duration appears unlikely. The analysis of correlation, however, provides arguments that cast doubts on the durational independence of the two vowels.

$V_1 \sim V_2$ correlation	S1	S2
Correlation coefficient	−0.14	−0.1
t test (n = 108)	−3.7	−2.74

Moreover, mean V_2 durations differ depending on the phonemic length of V_1, i.e. when V_1 is a phonemically short vowel, mean V_2 duration is always slightly greater than mean V_2 duration in items with a phonemically long/diphthongal V_1.

	S1	S2
Mean V_2 duration (short V_1)	143	204
Mean V_2 duration (long V_1)	141	199

One may argue that despite their statistical significance, the negative $V_1 \sim V_2$ correlations are rather weak and unconvincing. In our view this objection is unfounded and the result points at more than a chance regularity. It has to be remembered that while V_1 duration is naturally diversified due to phonemic length differences, PFC effects and intrinsic duration, V_2 is phonemically identical for all

[2] Note that the differences between V_1 mean duration in di- and trisyllables correspond roughly to the 30 ∼ 40 ms difference (S1 = 42 ms; S2 = 34 ms).

items as is its context (final open syllable). In such circumstances no correlation should be observed. It has to be admitted, though, that the differences in V_2 duration are unlikely to carry any perceptual load (cf. Lehiste 1970).

In trisyllabic items the conceivable set of temporal intervocalic relations is much larger than in disyllables and it includes the following possibilities:

- $V_1 \sim V_2$
- $V_1 \sim V_3$
- $V_2 \sim V_3$
- $(V_1 + V_2) \sim V_3$
- $(V_1 + V_3) \sim V_2$
- $(V_2 + V_3) \sim V_1$

It is only $V_1 \sim V_2$ interdependence that results in a statistically significant (negative) correlation (S1 = −0.11; *t* test = −2.98 and S2 = −0.25; *t* test = −7.20). The other durational relations are either statistically insignificant (for one or both subjects), statistically significant but of opposite value (positive vs. negative) or are a mixture of the two possibilities. The statistically significant and negative $V_1 \sim V_2$ correlation in trisyllables seems to be analogous to that observed between corresponding vowels in disyllables. The crucial difference, however, is that in disyllables V_2 was the word-final vowel, whereas in trisyllables it is the medial one. This may suggest that durational correlations between vowels are local, i.e. they involve vowels only in consecutive syllables. An alternative interpretation of the apparently non-systematic distribution of statistically significant intervocalic correlations in trisyllables is that individual speakers employ different networks of intervocalic durational correlations.

A fundamental problem, however, is why in the first place systematic and unsystematic (speaker-individual) significant correlations are observed. As far as qualitative interrelations between vowels in consecutive syllables (or within the entire word) are concerned, their explanation may be of articulatory nature, i.e. formant frequencies at the beginning of the following vowel are in a way 'inherited' from the formant frequencies observed in the final phase of a preceding vowel (Öhman Sven 1966) and vowel harmonies are related to a particular articulatory setting (nasality, openness/closeness). Generally, intervocalic qualitative relations may be considered 'spreading' or 'co-articulation' phenomena emanating from the stressed vowel. Intervocalic qualitative co-articulations, therefore, involve promoting a particular feature of the stressed vowel onto the unstressed ones within a domain. In terms of quantity, the only remote analogy we can think of is a simultaneous lengthening/shortening of all vowels within a domain connected with faster/slower tempo of delivery or phrase-final lengthening of the whole item. The durational correlations between the stressed and the unstressed vowels should then of necessity be positive. The significant $V_1 \sim V_2$ correlations observed in our data, however, are negative.

Given our experimental conditions, i.e. steady tempo of stimulus presentation, the fact that the increase in V_1 duration entails V_2 shortening suggests that there exists some pre-programmed durational pattern which controls the duration of both vowels

in disyllabic items. In other words, the duration of one vowel is checked against the duration of the other. This constitutes a serious argument against the 'no foot' hypothesis (e.g. Selkirk 1984) since it points to a superordinate temporal unit. Simultaneously, it accounts for why V_1 does not have to be longer than V_2 in disyllables or V_3 in trisyllables. The curtailed duration of V_1, e.g. when it is phonemically short and additionally affected by PFC, is thus 'compensated for' by the increased duration of V_2. If so, the durations of V_1 and V_2 in disyllabic items are both fine-tuned to fit a durational template whose overall duration, as we will argue in the next section, oscillates around 300 ms (similar results are reported by Kohno 1992).

In the following section additional arguments will be provided which support the assumption that intervocalic durational correlations are indeed superimposed by a higher-order durational template which not only enforces the equalisation of total vowel duration within items of the same number of syllables, but also levels off the durational differences between items having a different number of syllables.

3.3 Total Vowel Duration

As far as the degree of variation in total vowel duration within the groups of items with the same number of syllables is concerned, we observe that it is remarkably greater in the group of monosyllables than in polysyllables and in di- and trisyllabic words it is nearly identical. This also holds true for the standard deviation values.

	S1		S2	
	CoV	Std dev	CoV	Std dev.
Monosyllables	19.2	43.4	28.1	72.0
Disyllables	13.2	39.7	11.7	43.1
Trisyllables	13.9	38.3	12.7	43.0

Had there been no tendency to equalise the total vowel duration, an opposite regularity should be observed, i.e. the overall degree of variation in polysyllables, i.e. the summation of all individual vowel variations, should be higher than in monosyllables. Thus, the increase in the number of syllables should result in the increase of variation in total vowel duration, i.e. the more variables, the greater the variation. This, as we see, is not the case.

Another argument which directly supports the equalisation hypothesis is provided by the analysis of variation coefficients for individual vowels and their comparison with those obtained for total vowel duration in a particular group of items. In principle, the mean variation coefficient for the component vowels and for the total duration should be identical. The coefficients of variation (%) for individual vowels and total vowel duration are presented below.

(continued)

Disyllables	V_1	V_2		Mean		Total Vowel Duration
Disyllables	V_1	V_2		Mean		Total Vowel Duration
S1	12.9	12.1		18.0	>	13.2
S2	24.3	9.1		16.7	>	11.7
Trisyllables	V_1	V_2	V3	Mean		Total Vowel Duration
S1	28.9	22.7	10.9	20.8	>	13.9
S2	30.2	23.3	10.7	21.4	>	12.7

We observe instead that mean variation coefficients for individual vowels in di- and trisyllables are invariably greater than those obtained for total vowel duration and that for both subjects the difference between the mean variation coefficient for $V_1/V_2/(V_3)$ and total V duration CoV is slightly greater in tri- than in disyllables. Again, the increase in the number of variables (i.e. CoV of particular vowels) that may influence the total vowel duration is counterbalanced by the decrease in the degree of overall variability within each sample. The only motivation for this, somehow paradoxical, regularity seems to be the superimposed pressure on individual vowels to adjust their durations[3] in such a way that their accumulated duration fits a certain durational template. These facts, in our view, do point at a strong tendency towards the equalisation of total vowel duration within each group of items, which, as suggested above, directly refutes the 'no-foot' hypothesis. If, however, as we assume, the tendency has neural and aerodynamic foundations, then it should also, if not primarily, manifest itself in the equalisation of total vowel durations in items of different number of syllables.

Admittedly, when analysed in purely statistical terms, the differences in total vowel duration between 1-, 2- and 3-syllable words are significant ($p < 0.05$). However, the differences in mean total duration between di- and trisyllables are remarkably smaller than those between mono- and polysyllables. Interestingly, for both subjects the accumulated duration of vowels in disyllables is greater than that in trisyllables. This stands at variance with an intuition that the increase in the number of syllables within an item must entail the increase of its total vowel duration and indirectly supports the equalisation hypothesis.

Mean total vowel duration (ms)	S1	S2
Monosyllables	227.3	256.8
Disyllables	301.8	369.4
Trisyllables	274.2	338.7

[3] Understandably, the degree of the adjustment is restricted by a number of factors (vowel-intrinsic duration, phonemic length, consonantal context, minimal execution time, etc.).

In the light of previous research on perception of durational differences (e.g. Lehiste 1970), it is evident that the differences of the range observed for di- and trisyllabic items (S1 = 27.6 ms; S2 = 30.7 ms) are well below the level of perceptual significance.[4] For this reason, the actual statistical significance of these differences is of secondary importance. It is rather the relations between particular significances, mapped onto corresponding mean differences, that support the 'equalisation' hypothesis. Since inter-speaker variation in total vowel duration, however, does seem to be perceptually salient (>60 ms), isochrony is fundamentally a perceptual phenomenon (cf. Lehiste 1977), which nonetheless has its acoustic foundations. In other words, there exist non-reducible speaker-independent differences in total vowel duration (both within each group of items and between the two groups) which are mainly 'inherited' from the differences in stressed vowel duration (phonemic length, PFC effect). These differences, however, are counterbalanced by the variable duration of word-final vowels. In effect, the remaining discrepancy is neutralized perceptually. Thus, given the natural – often stylistically conditioned– variations in tempo in connected speech, the impression of rhythmicality is temporally local, i.e. it is confined to smaller parts of an utterance, e.g. tone units, as argued in Cauldwell (2002), or individual polysyllabic words, as in our experiment) and is then 'reset' before the following one begins. In effect, in a longer utterance there may be a few 'isochronies', corresponding to different parts of the same utterance. This interpretation is, on the one hand congruent with Cauldwell's postulate of 'functional irrythmicality' and, on the other, it explains why, despite stylistically conditioned and speaker-individual differences in tempo, the impression of isochrony is almost unanimously reported by listeners and consequently refuted by researchers (e.g. Roach 1982).[5]

3.4 Pitch-Duration Relations in Stressed and Unstressed Vowels

Apart from intervocalic quantitative relations we also analysed intravocalic pitch-duration interdependence. Although it may, at first sight, seem digressive, this interdependence is crucial for the interpretation of both the intervocalic durational correlations and the divergence of total vowel duration in di- and trisyllabic items.

[4] On the other hand, mean vowel duration in monosyllables is far greater than the minimal execution time (cf. Klatt 1976). This suggests that vowels in isolated and phrase final monosyllables undergo an extra lengthening, whose aim arguably is to accommodate a complex contour tone (Ciszewski, in prep.).

[5] Our study also shows that vowel-only approach to rhythm provides much less diversified results than the approach based on interstress interval. In our opinion, the latter has been wrong in assuming that both vowels and consonants contribute to the production/perception of rhythm (for more arguments see: Ciszewski 2010a).

Previous studies have shown that vowel duration (in acoustic and perceptual terms) is negatively correlated with mean f_0, i.e. low-tone vowels are longer than high-tone ones (Gandour 1977) and that dynamic tones require greater vowel duration (Gordon 2001; Zhang 2001; Yu 2002). A number of possible interpretations for this phenomenon are discussed in Ohala (1973). These include: (i) dynamogenetic theory by Taylor (1933), (ii) air pressure increase behind vowel constriction (Mohr 1971), (iii) vocal tract and vocal cords acoustic coupling (Atkinson and James 1972) and (iv) mechanical tongue-larynx interaction leading to vertical tensing of vocal folds (Ladefoged 1964). However interesting these studies may be, they all concentrate on duration-pitch interrelations in unreduced (hence stressed) vowels. Our results fully confirm earlier findings, but only as far as V_1 is concerned. Unstressed vowels seem to be subject to different laws of speech mechanics and aerodynamics.

In all three groups of target items (mono-, di- and trisyllables) a statistically significant negative correlation has been confirmed:

	S1		S2	
V_1 pitch-duration correlations	r	*t* test	r	*t* test
Monosyllables	−0.57	−5.97	−0.48	−20.46
Disyllables	−0.17	−4.60	−0.22	−7.59
Trisyllables	−0.49	−17.29	−0.18	−6.15

Since a systematic decrease in stressed vowel duration was independently observed as the number of the following unstressed syllables increases which is accompanied by a simultaneous increase in V_1 mean pitch, the negative pitch-duration correlation must, by inference, be much stronger if calculated for all groups of items collectively. This is indeed the case (S1: r = −0.67, *t* test = −96; S2: r = −0.65, *t* test = −89). Whichever theory is assumed to explain this regularity, its cross-linguistic validity does imply a mechanistic/aerodynamic motivation (which assumption the theories discussed by Ohala (1973) seem to share).

Any purely mechanistic or aerodynamic explanation, however, should cater simultaneously for stressed and unstressed vowels. What transpires from our data is that pitch and duration of unstressed vowels (both word-medial and word-final) are also coupled but the correlation is positive, rather than negative, as it was observed for stressed vowels.

	S1		*S2*	
V_1 pitch-duration correlations	r	*t* test	r	*t* test
Disyllables V_2 (final)	0.17	4.55	0.09	2.92
Trisyllables V_2	0.16	4.29	0.08	2.52
Trisyllables V_3 (final)	0.28	8.13	0.43	16.87

Thus, either the explanation is not mechanistic at all (which is unlikely) or some important factor(s?) has/have not been taken into consideration in earlier research.

3.5 Interpretation of Results

As we emphasised at the beginning of this section, the comparison of pitch-duration correlations in stressed and unstressed vowels is not at all digressive and it is directly related to the systematic intervocalic quantitative (negative) $V_1 \sim V_2$ correlation and the equalisation of total vowel duration. Only when these types of observations are 'mapped' onto each other, can the –somewhat paradoxical– difference in pitch-duration correlation between stressed and unstressed vowels be fully understood.

Our explanation is as follows. The onset of the stressed vowel coincides with a sudden increase of subglottal pressure (cf. Ladefoged 1967: 46) produced by appropriate muscular constrictions, which in turn results in the increase of short time average volume velocity. Due to Bernoulli's effect the relation between the velocity of air passing through the glottis and time is inversely proportionate. Thus, the longer the time between the outburst of acoustic energy (which is a function of volume velocity and subglottal pressure) and the complete occlusion caused by the following consonant, the greater the decrease in velocity, and consequently, the average pitch of the stressed vowel (= negative pitch-duration correlation in stressed syllables). This mechanism is also responsible for V_1 pitch increase in di- and trisyllables (note that V_1 duration is inversely proportionate to the number of the following unstressed syllables) and explains why V_1 pitch slope is significantly greater in monosyllables than in polysyllables (the greater distance between the initial energy outburst and consonant occlusion does not allow to maintain stable air velocity; the steady decrease in velocity results in the simultaneous decrease of f_0) (Fig. 3).

Unstressed vowels, on the other hand, are not accompanied by the increase of subglottal pressure. They 'inherit' their energy from the preceding vowel ($V_1 > V_2 > V_3$).[6] Since subglottal pressure does not increase, air velocity is relatively stable. Depending on the amount of energy which remains after the articulation of the stressed vowel (note that it may vary due to the differences in phonemic length and/or the operation of PFC), the following unstressed vowels may also vary in duration. Thus, if V_1 is shorter, it has higher pitch; this in turn elevates V_2 (V_3) pitch proportionately. For the same reason, V_2 is proportionately longer (= negative durational $V_1 \sim V_2$ correlation). This is reflected in the positive pitch-duration correlation in unstressed syllables.

[6] Similarly to the terminal frequencies of formants in VCV utterances which have been shown to depend on the entire vowel context (Öhman 1966).

Although both the quantitative intervocalic relations and intravocalic pitch-duration relations may be interpreted mechanically and aerodynamically, they may also perform important communicative functions. In particular we hypothesise that it is the acoustic characteristics of V_1 that facilitates word processing and recognition. For instance, relatively low pitch, accompanied by substantial pitch slope and greater duration signals the proximity of a word boundary. In contrast, higher but level pitch together with reduced duration and greater intensity signal a more remote word boundary.

	Duration	Pitch (mean	Pitch slope	Perceptual message
V_1 *bid*	241 ms	102 Hz	41 Hz	THE END
V_1 *biddy*	133 ms	125 Hz	15 Hz	1 SYLL. TO THE END
V_1 *bigamy*	91 ms	135 Hz	7 Hz	2 SYLL. TO THE END

Thus, the relations and regularities which first of all are mechanistically/aerodynamically conditioned have an independent perceptual value.

4 Conclusions

The results obtain in our experiment suggest that intervocalic relations in morphologically simple, initially stressed lexical items are not only qualitative, but also quantitative. It has been found that the durations of stressed and post-stress vowels are bound by statistically significant negative correlation. This indicates that the vowels in question 'negotiate' their durations, i.e. the increase in the duration of V_1 entails the decrease in V_2 duration. The fact that the quantitative relation is negative may be adequately explained only if we assume that it is controlled by a superimposed durational template. Thus, the durational adjustments aim at equalising the total vowel duration not only within items having the same number of syllables (but differing in V_1 duration) but also in di- and trisyllabic words. The equalisation is additionally supported by statistically significant differences in final vowel duration, which suggests that the degree of final lengthening is also controlled by the same durational template. We hypothesise that this template corresponds to stress foot. The vowel-only approach, which was assumed in this study, provides strong arguments in favour of the isochrony hypothesis which has been consequently refuted for interstress intervals which comprise consonants as well. The unit of pre-programmed timing which emerges from our data corresponds to approximately 300 ms.[7]

[7] Admittedly, this figure may be dependent on speech rate and an analysis of spontaneous speech is likely to reveal a much greater durational variability.

Apart from the quantitative intervocalic relations we also observed intriguing intravocalic correlations between duration and pitch. The analysis confirms earlier findings which point at a cross-linguistically valid negative correlation between the two correlates of stress. This correlation, however, holds only for stressed vowels. Unstressed vowels display an opposite pitch-duration correlation.

Both the quantitative intervocalic correlations and the variable pitch-duration correlations in stressed and unstressed vowels are, in our view, mechanistically and aerodynamically conditioned. Their overall acoustic effect, i.e. a combination of V_1 duration, equalised total vowel duration, V_1 pitch increase in longer items and the remarkably greater pitch and intensity slope in monosyllables, however, are perceptually informative. They aid word processing and lexical access (as suggested by McAllister 1991, van Donselaar, Koster and Cutler 2005, among others) and serve as prosodic boundary distance markers.

Further research on perception of these phenomena is required to verify our findings.

References

Atkinson, James E. 1972. Intrinsic F0 in vowels: Physiological correlates. Paper presented at Fall ASA meeting, Miami Beach

Boersma, Paul and David Weenink. 2005. Praat: doing phonetics by computer (version 5.0.32). http://www.praat.org/ (13 August 2008)

Cauldwell, Richard. 2002. The functional irrhythmicality of spontaneous speech: A discourse view of speech rhythm. Applied Language Studies, Apples 2/1: 1–24

Ciszewski, Tomasz. 2010a. Theoretical modelling of word-level isochrony in English. In *Issues in Accents of English* 2: *Variability and Norm*: 261–275. Cambridge Scholars Publishing

Ciszewski, Tomasz. 2010b. Is metrical foot a phonetic object? In *Research in Language*, Waniek-Klimczak, E. & L. Shockey (eds.), Łódź University Press: 115–135

Ciszewski, Tomasz (in prep.) *The anatomy of the English metrical foot: acoustics, perception and structure*

Fabricius, Anne. 2002. Weak vowels in modern RP: An acoustic study of happY-tensing and KIT/schwa shift. *Language Variation and Change* 14:211–237. Cambridge University Press

Fletcher, Janet. 2010. The prosody of Speech: Timing and Rhythm. In Hardcastle, W.J., Laver, F. and F. E. Gibbon (eds.), *The Handbook of Phonetic Sciences.* (Second edition). Blackwell Publishing Ltd

Gandour, Jackson. 1977. On the interaction between tone and vowel length: Evidence from Thai dialects. Phonetica 3:54–65

Gordon, Matthew 2001. A typology of contour tone restriction. Studies in Language 25(3):423–462

Klatt, Dennis H. 1976. Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. Journal of the Acoustical Society of America 59:1208–1221

Kohno, Morio. 1992. Two mechanisms of processing sound sentences. In Tohkura, Y., Vatikiotis-Bateson, E. and Y. Sagisaka (eds.), *Speech Perception, Production and Linguistic Structure*: 287–29. Amsterdam, Washington, Oxford: IOS Press

Ladefoged, Peter. 1964. Some possibilities in speech synthesis. Language and Speech 7:205–14

Ladefoged, Peter. 1967. *Three areas of experimental phonetics.* Oxford University Press

Lehiste, Ilse. 1970. Suprasegmentals. Cambridge, MA: The MIT Press

Lehiste, Ilse. 1977. Isochrony reconsidered. Journal of Phonetics 5: 193–-212

McAllister, Jan. 1991. The Processing of Lexically Stressed Syllables in Read and Spontaneous Speech. *Language and Speech January 1*, 34:1–26

Mohr, Buurckhard. 1971. Intrinsic variations in the speech signal. Phonetica 23: 65–93.

Ohala, John. 1973. Explanations for the intrinsic pitch of vowels. Unpublished manuscript

Öhman Sven E. G. 1966. Coarticulation in VCV utterances: Spectrographic measures. *Journal of the Acoustical Society of America*, 39: 151-168.

Roach, Peter. 1982. On the Distinction between "Stressed Timed" and "Syllable Timed" Languages. In Crystal, D. (ed.): *Linguistic Controversies. Essays in Linguistic Theory and Practice.* In Honour of F. R. Palmer: 73–79.

Selkirk, Elisabeth. 1984. *Phonology and Syntax. The relation between sound and structure.* Cambridge, MA: MIT Press

Taylor, Herbert. C. 1933. The fundamental pitch of English vowels. *Journal of Experimental Psychology* 16: 565–582.

van der Hulst, Harry and Jeroen van de Weijer. 1995. Vowel Harmony. In Goldsmith, J.A. (ed.): *Handbook of Phonological Theory*, Cambridge, MA: Blackwell Publishers

van Donselaar, Wilma, Koster, Mariëtte & Cutler, Anne. (2005). Exploring the role of lexical stress in lexical recognition. Quarterly Journal of Experimental Psychology, 58A(2): 251–273

Waniek-Klimczak, Ewa. 2005. *Temporal parameters in second language speech. Łódź: Wydawnictwo Uniwersytetu Łódzkiego*

Yip, Moira. 1995. Tone in East Asian Languages. In Goldsmith, J.A. (ed.): Handbook of Phonological Theory, Cambridge, MA: Blackwell Publishers

Zhang, Jiajie. 2001. *The effects of duration and sonority on contour tone distribution-typological survey and formal analysis*. PhD dissertation. Department of Linguistics. University of California, Los Angeles

Segmental and Prosodic Reasons for Non-Native Timing of *her* in Polish Learners, English Speech

Andrzej Porzuczek

Abstract Foreign speech timing problems depend on both segmental and prosodic factors. Thus inadequate durations and duration proportions may result from segmental substitutions involving sounds of different intrinsic length, activation of L1 instead of FL sandhi processes and generally non-native temporal organisation of utterances caused by the use of L1 rhythmic patterns. This paper focuses on absolute duration and timing relations of the unstressed pronoun *her* pronounced by a group of Polish learners in three phrases appearing in a read text: (1) but everyone called *her* Cinders. (2) When *her* sisters had gone, (3) It was *her* fairy godmother. The respondents, thirteen college students of English, were recorded at the beginning of their first year and seven months later, after two semesters of general English and practical phonetic training. The data were then compared to native Southern British English production recorded in the IViE database. Discrepancies in absolute and relative duration of the pronoun are believed to be connected with the substitution of /x/ for /h/, a consonant prone to elision in native English speech, insufficient vowel reduction, and rhoticity of the coda. The data show the scale of the problem and Polish learners' development in the course of practical phonetic training. The first recording reveals significantly longer durations in Polish learners' performance, magnified by fluency problems. The second recording, preceded by pronunciation training, indicates smaller length differences between Polish and English speakers. It also turns out that the duration of onset in Polish learners' performance contributes more to non-native timing of the weak form than its rhyme.

Research supported by a grant from Polish Ministry of Science and Higher Education (0576/B/H03/2010/38).

A. Porzuczek (✉)
University of Silesia, Katowice, Poland
e-mail: ampj24@wp.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_13,

1 Introduction

Foreign language (FL) speech timing problems may contribute not only to foreign-accentedness of the learner's pronunciation but also to communication failures. The importance of timing for intelligibility has been confirmed by numerous studies, e.g. Kozhevnikov et al. (1965), Wingfield and Klein (1971), Jassem (1971), Adams (1979), Faber (1986), Tajima et al. (1997), Smith et al. (2003). Therefore Polish learners of English should be aware of temporal discrepancies between the two languages in order to avoid negative transfer. The phonologically relevant features connected with timing which are often neglected by Polish learners include the following:

- intrinsic vowel length distinction (Sobkowiak 1996; Szpyra-Kozłowska 2003; Nowacka 2008; Bryła 2010),
- vowel length as a cue to underlying coda voicing (Sobkowiak 1996; Szpyra-Kozłowska 2003; Waniek-Klimczak 2005; Nowacka 2008),
- VOT as a cue to underlying onset voicing (Waniek-Klimczak 2005; Rojczyk 2008).

Apart from the factors mentioned above, timing depends on the distribution of prominence in an utterance. Although stressed speech units are universally longer than unstressed ones, durational contrast between them is smaller in Polish, which results in insufficient unstressed syllable reduction and too short prominent syllables (Avery and Ehrlich 1992; Hewings 2004; Dziubalska-Kołaczyk et al. 2006; Gonet et al. 2010).

Finally, the mean length of a prosodic unit realised by the learners tends to differ significantly from native scores even if only some learners fail to activate typical FL lenition processes, especially elision. Natural L1 elision processes are often suppressed in the pronunciation of foreign learners, who are strongly influenced by spelling (cf. Luke et al. 1982; Sobkowiak 1996; Hewings 2004; Gonet et al. 2010). Needless to say, *her* is a word where as many as two superfluous segments can appear in learner speech. This is an important reason for presenting individual subjects' scores in the following sections of this paper.

2 Objectives of the Study and Research Procedure

This study focuses on the duration of the pronoun *her* in unstressed positions in the read speech of Polish learners and native English speakers. The word has been selected for observation because, if analysed both in terms of absolute duration and in relation to the other phrase components, it can reflect most of the non-native timing problems mentioned in the introduction. Its phonological structure and graphic representation may cause Polish learners to produce a relatively long syllable if they.

- substitute a Polish velar fricative for the glottal sound (frequently elided by native speakers) in the onset
- fail to reduce the vowel,
- articulate the final /r/, typically using vowel and consonant in linear succession rather than retroflection
- fail to proportionally strengthen the prominent elements in the unit.

This paper focuses on the timing of the pronoun and its constituents, while acknowledging the necessity to supplement duration studies with spectral analysis of the vowel (including the issue of rhoticity) and the onset consonant, and correlate it with F0 and intensity cues.

The word and its constituent segments have been measured in Polish learners' read text, a part of *Cinderella* story, borrowed from the IViE corpus (Grabe et al. 2001, see Appendix). Three occurrences of the word have been analysed in the following phrases, which show highly uniform prominence distibution in both groups of speakers:

1. but everyone called *her* Cinders.
2. When *her* sisters had gone, …
3. It was *her* fairy godmother.

The respondents were 13 college students of English at the beginning of their first year (POL1). They had been given time for preparation prior to the reading, including the opportunity to consult the pronunciation of problematic words. The recording took place in quiet conditions of the college's language laboratory. We used a Panasonic RR-US360 portable digital recorder, placed at a comfortable distance of approximately 25 cm from the respondent's mouth. The signal was sampled at 16 kHz and transferred to a computer's hard disk in the wav format for further analysis by means of the Praat software package 4.4.30 (Boersma 2001). The same recording procedure was repeated after 2 semesters (7 months) of general English and standard practical phonetic training (POL2). The data were then compared to native Southern British English production of 12 secondary school students in Cambridge, recorded in the IViE corpus.

The word *her* was measured as a sum of its onset and rhyme (/h/+/er/). The beginning of /h/ in phrases (1) and (2) was marked by the release of oral closure for /d/ and /n/. In phrase (3) it was the offset of high frequency noise corresponding to alveolar friction. The rhyme was measured from the onset of clear vocalic formant structure to the onset of frication marking the beginning of the following syllable. The rhyme was not further divided into segments in cases of rhotic pronunciation.

The duration of *her* was also compared to the duration of the whole measured text portion. Each of the three tested phrases contained two accented syllables, which allowed us to observe timing proportions between prominent and non-prominent components. Investigating the rhythmic properties of speech requires relative timing measures, although it must be remembered that absolute length data are also important since timing relations may vary at different speech rates (e.g. Dauer 1983; Dellwo and Wagner 2003; Dellwo 2008).

Table 1 Mean absolute (ms) and relative (% phrase) duration of *her* in …*called her Cinders* (c-her), *When her sisters had gone* (w-her), and *It was her fairy godmother* (iw-her)

Ms↓	Mean (SD)	c-her	w-her	iw-her	%phrase	c-her	w-her	iw-her
ENG	**97 (25.2)**	117	86	89	ENG	15.9 %	8 %	7.6 %
POL2	**127 (54.5)**	128	116	136	POL2	16.7 %	9.5 %	9.7 %
POL1	**196 (68.4)**	223	157	209	POL1	24.4 %	11.9 %	12.7 %
POL:ENG ratio ↓	↓	↓	↓	↓	POL:ENG ratio ↓	↓	↓	↓
POL2:ENG	**1.31**	1.1	1.36	1.53	POL2:ENG	1.05	1.19	1.28
POL1:ENG	**2.02**	1.9	1.83	2.36	POL1:ENG	1.53	1.49	1.67

The initial part of the first phrase (*but everyone*) was not measured because in several cases *everybody* was substituted for *everyone*, which inevitably affected the ratios. Moreover, the final /z/ in *Cinders* was not measured either, because it is hard to establish the moment when the aperiodic signal intensity drops below the threshold of perception. Such modifications of the context do not influence the results as long as we analyse the relations between two groups of subjects rather than try to establish modal durations of utterances and their constituents.

3 General Results

Since the length of *her* has been found to vary across the contexts, the results illustrating mean absolute (ms) and relative (% phrase) durations of the pronoun and its parts are presented in Table 1 en masse and separately for the three individual phrases.

Fairly advanced Polish learners (POL1) tend to make the pronoun twice as long (102 %) as the native speakers in terms of absolute duration and over 50 % longer in terms of timing proportions within the measured phrase. Increased English learning experience, including practical phonetic training, reduces the difference to approximately 30 % in absolute duration (see POL2) and 5–28 % depending on the context in timing proportions. A review of the results for individual speakers shows considerable variation, suggesting diverse abilities of individuals to acquire native-like timing. Individual scores for all three contexts are shown in Fig. 1.

Apart from one outlier, all native English samples are between 50 and 150 ms long. Many POL2 (approximately a half) and some POL1 responses are confined within this range too, but in numerous cases, especially in the first recording session, Polish performance is considerably longer, with 9 responses over 250 ms. On the one hand, we recognise general fluency problems, including more frequent stops and hesitations as the main source of these very long durations, but on the other, even in the fluent responses the length of *her* hardly approaches native values, which means that dysfluencies are not the only problem.

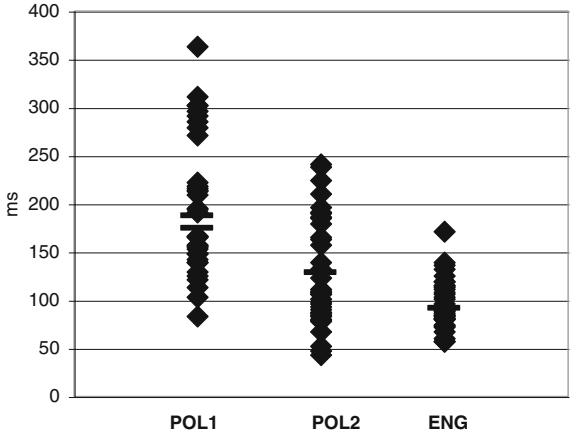


Fig. 1 Individual durations of *her*. Horizontal bars indicate group medians

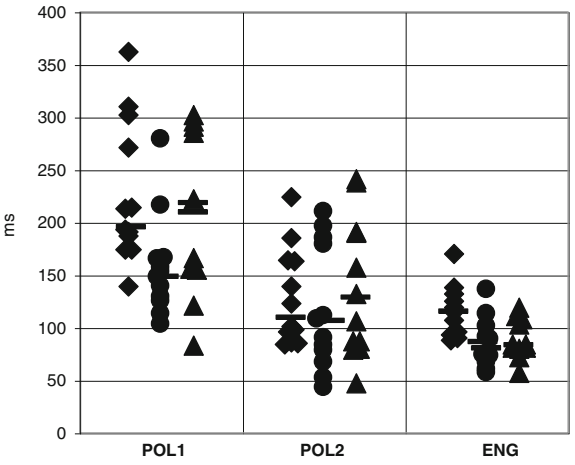


Fig. 2 The duration of unstressed *her* in…*called her Cinders* (*squares*), *When her sisters had gone* (*circles*), and *It was her fairy godmother* (*triangles*). Horizontal bars indicate group medians

It is also worth noting that the native English speakers make the duration of *her* longer in …*called her Cinders* than in the other two contexts, where no significant differences have occurred. No such pattern has been observed in Polish learners' production. As a matter of fact, in both recording sessions, the learners' median durations are the longest in *It was her fairy godmother*, where the native speakers make the word shorter than in the other phrases.

Finally, individual variation is much smaller in the pronunciation of native speakers', which reflects more articulatory consistency of that group. The distribution of responses related to each of the three contexts is shown in Fig. 2.

A statistical test shows (Table 2) large initial discrepancy in the duration of *her* between the two groups (POL1:ENG), and more native-like values for POL2.

Training allows the learners to achieve native-like durations of *her* in *called her*. The progress is not obviously significant in *When her*, where the initial difference

Table 2 One-way ANOVA p-values for between-group difference in the duration of *her*

her in	POL1:POL2	POL1:ENG	POL2:ENG
called ___	<.001	<.001	**.422**
when ___	.056	<.001	**.099**
it was___	.014	<.001	.021

between native and non-native durations is smaller, but it is enough to reduce the cross-group discrepancy to a statistically non-significant level. Finally, in *It was her…*, the reduction is statistically significant, but still seems to be more difficult for the learners than in *called her*, and the POL2:ENG difference remains considerable.

4 Segmental Duration

The next phase of the analysis shows the contribution of individual segments to the total duration of the word. Figure 3 presents the duration of the consonantal onset and the syllable rhyme in the three tested contexts.

The duration of /h/ in POL1 reveals a significant problem with consonant reduction. However, a native-like mean length of the segment is achieved after 7 months of practice mainly by dropping the initial /h/, widespread in native English pronunciation of weak forms, but also by shortening the consonant if it is not elided. The low prominence level of the function word is also reflected in rhyme duration reduction, which typically results in a native unstressed vowel 50–100 ms long. In their first reading, the Polish learners use such a short vowel only in *when her sisters*, while even the medians for the other two contexts far exceed the 100 ms upper limit set by the native speakers. The scores, however, become much closer to the native performance in the second recording.

These results suggest that vowel duration is not the only problem with unstressed syllable reduction. An analysis of variance actually indicates a higher level of significance of the group factor for the duration of onset than rhyme (Table 3).

The results for individual contexts in paired group relations are shown in Table 4.

"h" is reduced more radically than "er" by the learners in the second recording, which practically reaches native standards in all contexts but *It was her.…*This is connected with a stronger tendency in native speakers to reduce /h/ after fricatives than in other contexts (cf. Shockey 2003:44). In the Polish learners' performance, /h/-reduction is similar in all three positions.

Vowel reduction is slighter (in *when her* actually non-existent) but generally enough to approach native-like durations. Naturally, given that the elision of /h/ can dramatically influence duration means, we cannot draw far-reaching conclusions, but the problem of initial /h/ articulation should not pass unnoticed in English pronunciation teaching to Poles.

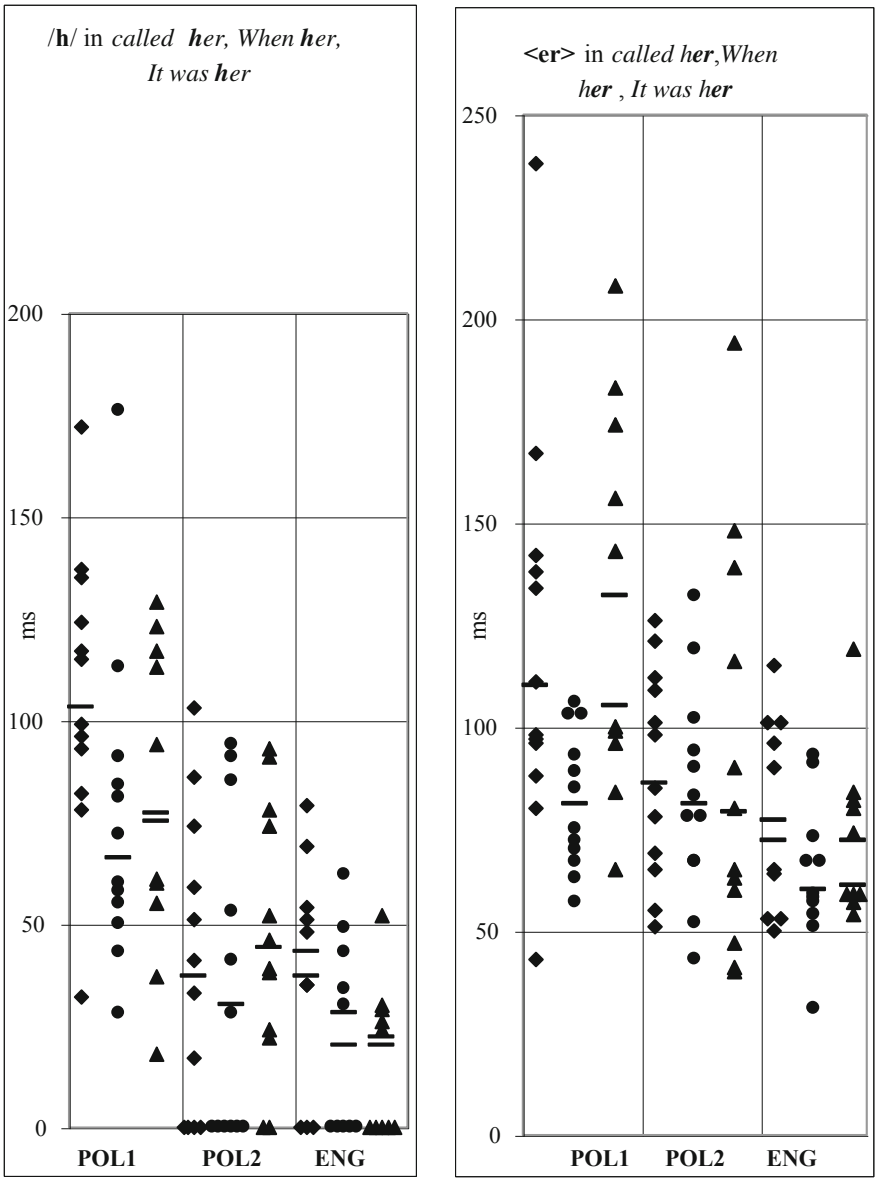


Fig. 3 The duration of onset and rhyme in individual speakers' pronunciation of *her* in…*called her Cinders* (*squares*), *When her sisters had gone* (*circles*), and *It was her fairy godmother* (*triangles*). Horizontal bars indicate group medians

Table 3 One-way ANOVA F-values for overall between-group difference

Segment\phrase	Called **her F (2,35)**	When **her F (2,35)**	It was **her F (2,34)**
h	19.36 ($p < .001$)	8.97 ($p < .001$)	14.15 ($p < .001$)
er	4.87 ($p = .014$)	3.94 ($p = .029$)	6.51 ($p = .004$)

Table 4 One-way ANOVA p-values for between-group difference in "h" and "er" duration

h in	POL1:POL2	POL1:ENG	POL2:ENG
called ___	<.001	<.001	**.938**
when ___	.008	<.001	**.417**
it was___	.019	<.001	.008
er in	POL1:POL2	POL1:ENG	POL2:ENG
called ___	.058	.013	**.259**
when ___	**.838**	.011	.028
it was___	.042	<.001	**.234**

Table 5 Mean unit duration (ms) and timing proportions in…*called her Cinder(s)*

Group\unit	Called	**her**	Cind	er	Σ
ENG (ms)	183	**117**	313	123	737
POL2 (ms)	202	**128**	296	142	769
POL2/ENG	**1.1**	**1.1**	**.95**	**1.16**	**1.04**
POL1 (ms)	221	**223**	318	152	914
POL1/ENG	**1.2**	**1.9**	**1.02**	**1.23**	**1.24**

5 Phrase-Level Timing Relations

So far we have analysed the absolute durations of the pronoun *her* and its segmental components and compared the learners' performance to native pronunciation models. Since we hypothesise that the main reasons for the discrepancies lie in the prosodic characteristics of native English, especially the contrast between prominent and non-prominent phrase units, we must make sure that the differences found in Polish learners' speech are not due to articulatory problems with foreign sounds or generally slower speech rate. For this reason, we have analysed length relations between phrase elements. Table 1 showed a higher percentage of total phrase duration contributed by the articulation of *her* for Polish speakers in all three contexts. Tables 5, 6, 7 display more detailed results.

In each case the accented syllables or words (*called-Cind, sisters-gone, fairy-god*) are comparable in length across the groups, especially the nuclear syllables *Cind*, *gone* and *god*, where the difference never reaches 10 %, even in POL1. On the other hand, the pronoun *her* and the other unstressed function words (*had, it, was, her*) are consistently longer in both Polish speakers' recordings. A smaller difference in phrase-final unstressed syllables suggests a more pronounced final lengthening effect in native English speakers.

Table 6 Mean unit duration (ms) and timing proportions in *When her sisters had gone*

Group\unit	when	**her**	sisters	had	gone	Σ
ENG (ms)	124	**86**	418	119	319	1065
POL2 (ms)	160	**116**	412	181	325	1194
POL2/ENG	**1.29**	**1.36**	**.98**	**1.53**	**1.02**	**1.12**
POL1 (ms)	148	**157**	453	210	340	1309
POL1/ENG	**1.2**	**1.83**	**1.08**	**1.77**	**1.07**	**1.23**

Table 7 Mean unit duration (ms) and timing proportions in *It was her fairy godmother*

Group\unit	it	was	**her**	fairy	god	mo	ther	Σ
ENG (ms)	66	143	**89**	341	249	121	165	1173
POL2 (ms)	125	189	**136**	347	270	173	163	1402
POL2/ENG	**1.9**	**1.32**	**1.53**	**1.02**	**1.09**	**1.43**	**.99**	**1.2**
POL1 (ms)	132	221	**209**	453	267	186	175	1642
POL1/ENG	**2**	**1.54**	**2.36**	**1.33**	**1.07**	**1.54**	**1.06**	**1.4**

Table 8 P-values for a paired *t* test for POL1:POL2 durational difference in *her, h-* and *–er*

Unit\context	called ___	when ___	it was___
her	<.001	.034	.003
h	<.001	.009	.003
er	.015	**.424**	.019

6 Individual Polish Learner's Progress

In order to establish whether the shorter durations in POL2 indicate significant progress of individual learners, a paired *t* test has been performed (Table 8).

The results show a significant shortening of all the measured units in the performance of the individual speaker, with the exception of the vowel in *When h**er***, where the initial difference between Polish learners and native English speakers was the least significant.

7 Conclusions

The results of the study confirm the common belief that Polish learners face problems with weak form reduction. In the case of *her*, the durational difference between Polish learners and native English speakers exceeds 100 % in pre-training performance (Table 1) and was in this case reduced to 31 % after the training. It appears that vowel reduction is not the main reason for weak-form shortening failures, even if it may be combined with additional rhoticisation, A reduction of /h/, observed in many learners after the training, helps achieve more native-like durations of *her*.

The length of the tested word in native production also depends on the prosodic and/or syntactic position in the phrase. This study does not provide enough evidence to draw firm conclusions but a tentative claim that native English speakers may reduce possessive adjectives more than personal pronouns in the objective case (cf. *called her Cinders* vs *When her sisters…*) provides a hypothesis for further research. In Polish learners, however, no significant difference between these contexts has been observed. On the other hand, weak-form reduction in Polish learners is less successful in the prehead, where native speakers use the shortest forms (cf. Jassem 1952). Alternatively, the learners' problems might be due to a longer sequence of unstressed syllables in *It was her…* in comparison with trochaic *When her sisters…* and…*called her Cinders*.

Finally, more language learning experience, including practical phonetics training, leads to smaller discrepancies between the learners' interlanguage and native English timing norms. Instruction and practice helps the students reduce the consonant more efficiently than the vowel but it is not clear if the reason for the onset-rhyme progress differences is greater initial cross-group discrepancy in onset duration or higher difficulty level of vowel reduction.

A.1 8 Appendix

Test Passage

Once upon a time there was a girl called Cinderella. But everyone **called her Cinders**. Cinders lived with her mother and two stepsisters called Lily and Rosa. Lily and Rosa were very unfriendly and they were lazy girls. They spent all their time buying new clothes and going to parties. Poor Cinders had to wear all their old hand-me-downs! And she had to do the cleaning!

One day, a royal messenger came to announce a ball. The ball would be held at the Royal Palace, in honour of the Queen's only son, Prince William. Lily and Rosa thought this was divine. Prince William was gorgeous, and he was looking for a bride! They dreamed of wedding bells!

When the evening of the ball arrived, Cinders had to help her sisters get ready. They were in a bad mood. They'd wanted to buy some new gowns, but their mother said that they had enough gowns. So they started shouting at Cinders. 'Find my jewels!' yelled one. 'Find my hat!' howled the other. They wanted hairbrushes, hairpins and hair spray.

When her sisters had gone, Cinders felt very down, and she cried. Suddenly, a voice said: 'Why are you crying, my dear?'. **It was her fairy godmother**!

References

Adams, Corinne. 1979. English speech rhythm and the foreign learner. The Hague: Mouton.
Avery, Peter; Ehrlich, Susan. 1992. Teaching American English Pronunciation. Oxford: Oxford University Press.

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. Glot International 5 (9/10):341–345.

Bryła, Agnieszka. 2010. Phonetic properties of Euro-English – empirical evidence. In: Waniek-Klimczak, Ewa (ed.). Issues in accents of English 2: Variability and norm. Newcastle-upon-Tyne: Cambridge Scholars Publishing, 37–60.

Dauer, Renate. 1983. Stress timing and syllable timing reanalyzed. Journal of Phonetics 11:51–62.

Dellwo, Volker. 2008. The role of speech rate in perceiving speech rhythm. *Speech Prosody 2008*, Proceedings of Speech Prosody 2008 series. Campinas, 375–378

Dellwo, Volker; Wagner, Petra. 2003. Relationships between speech rhythm and rate. *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, 471–474

Dziubalska-Kołaczyk, Katarzyna; Bogacka, Anna; Pietrala, Dawid; Wypych, Mikołaj; Krynicki, Grzegorz. 2006. PELT: an English language tutorial system for Polish speakers. *MULTILING-2006*, paper 012

Faber, David. 1986. Teaching the rhythms of English: A new theoretical base. International Review of Applied Linguisticsin Language Teaching 24:205–216.

Gonet, Wiktor; Szpyra-Kozłowska, Jolanta; Święciński, Radosław. 2010. The acquisition of Vowel Reduction by Polish students of English. In: Waniek-Klimczak, Ewa (ed.) *Issues in Accents of English 2: Variability and Norm.* Newcastle-upon-Tyne: Cambridge Scholars Publishers, 291–308

Grabe, Esther; Post, Brechtje; Nolan, Francis. 2001. *The IViE Corpus*. Department of Linguistics, University of Cambridge. http://www.phon.ox.ac.uk/IViE. [Retrieved 7 September 2006]

Hewings, Martin. 2004. Pronunciation Practice Activities. Cambridge: Cambridge University Press.

Jassem, Wiktor. 1952. *Intonation of Conversational English (Educated Southern British).* Wrocław: Wrocławskie Towarzystwo Naukowe.

Jassem, Wiktor. 1971. Podręcznik wymowy angielskiej. Warszawa: Państwowe Wydawnictwo Naukowe.

Kozhevnikov, Valery; Chistovich, Ludmila; 1965. *Speech: Articulation and Perception*. Washington, D. C.: Joint Publication Research Service No. 30,543. (Translated from *Rech: Artikulyatsiya iVospriyatiye*. Moscow-Leningrad, 1965)

Luke, Kang Kwong; Richards, Jack C. 1982. English in Hong-Kong: Functions and status. *English World-Wide* 3:147–164.

Nowacka, Marta. 2008. *The Phonetic Attainment in Polish University and College Students of English. A Study in the Productive and ReceptivePronunciation Skills.* Unpublished Ph.D. dissertation. Maria Curie-Skłodowska University, Lublin

Rojczyk, Arkadiusz. 2008. *Perception of English and Polish Obstruents*. Unpublished PhD dissertation. University of Silesia, Katowice.

Shockey, Linda. 2003. Sound Patterns of Spoken English. Oxford: Blackwell Publishing

Smith, Bruce; Bradlow, Ann R.; Bent, Tessa. 2003. Production and perception of temporal contrasts in foreign accented English. In: Sole, Maria-Josep; Recasens, Daniel; Romero Joaquin (eds.). *Proceedings of the XVth International Congress of Phonetic Sciences*. Barcelona, 519–522

Sobkowiak, Włodzimierz. 1996. English Phonetics for Poles. Poznań: Bene Nati

Szpyra-Kozłowska, Jolanta. 2003. The Lingua Franca Core and the Polish Learner. In: Sobkowiak, Włodzimierz; Waniek-Klimczak, Ewa. (eds.). *Dydaktyka fonetyki języka obcego*. Płock: Wydawnictwo Naukowe PWSZ w Płocku, 193–210

Tajima, Keiichi; Port, Robert; Dalby, Jonathan. 1997. Effects of temporal correction on intelligibility of foreign-accented English. Journal of Phonetics 25:10–24.

Waniek-Klimczak, Ewa. 2005. *Temporal Parameters in Second Language Speech.* Łódź: Wydawnictwo Uniwersytetu Łódzkiego

Wingfield, Arthur; Klein, Jan F. 1971. Syntactic Structure and Acoustic Pattern in Speech Perception. *Perception in Psychophysics* 9:23–25

Part IV
Sociolinguistics

Stratification of English as a Lingua Franca: Identity Constructions of Learners and Speakers

Višnja Josipović Smojver and Mateusz-Milan Stanojević

Abstract Studies of English as a Lingua Franca (ELF) suggest that some speakers of ELF are willing to mark their (national) identity in their ELF pronunciation, which we call the liberal stance, while others want to strive towards native models (e.g. Jenkins, World Englishes 28:200–207, 2009), which we describe as the conservative stance. A recent study (Stanojević and Josipović, Euro-English and Croatian national identity. Conference paper presented at the New Challenges for Multilingualism in Europe, Dubrovnik, Croatia, April 11–15, 2010) confirms this, suggesting that liberalism versus conservativism towards ELF among Croatian university students is correlated with their major field of study. In this paper we show that there may be a more pervasive process at play behind the liberal versus conservative attitudes to ELF, namely identity construction. Based on the results of a questionnaire conducted among secondary school pupils, university students and employees of a company, we show that different attitudes to one's own accent, the accent of one's conversational partners and teaching models primarily hinge on belonging to different groups of participants: learners versus speakers of ELF. Our results support a non-monolithic, stratified ELF model, which allows changes in accordance with the needs and identity construction of its speakers.

V. J. Smojver (✉) · M.-M. Stanojević
Faculty of Humanities and Social Sciences, University of Zagreb, Zagreb, Croatia
e-mail: visnja.josipovic@zg.t-com.hr

M.-M. Stanojević
e-mail: mmstanoje@ffzg.hr

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_14,

1 Introduction

The fastest-growing function of English in the world is as a Lingua Franca (ELF), i.e., as the common language in the interaction among non-native speakers from different linguistic and cultural backgrounds (cf. Görlach 2002; Mauranen and Ranta 2009). It has often been observed that the speakers of this variety (or class of varieties) of English definitely outnumber native speakers. As a result, an increasing number of prominent linguists have begun to specialise in various aspects of ELF. Probably the best known among them are those dealing with the nature and function of ELF as the only truly global language, e.g., works by Crystal (2004), Jenkins (2006a, 2007, 2009), Seidlhofer (2001) or Dewey (2007); or those about suggested directions for research (cf. Seidlhofer 2009). Likewise, in recent years linguists have also been concerned with the identification and description of the various communities of practice for ELF, thus describing subvarieties of ELF such as business ELF, also known as BELF, e.g. Ehrenreich (2009) and Pullin Stark (2009); or ELF as an academic language or, similarly, ELF in classroom communities of practice, Smit (2009). There have also been some widely quoted works on ELF variability on all levels of linguistic analysis, such as syntax, e.g. in Dewey (2009) or Ranta (2009); semantics and phraseology, e.g., Pitzl (2009); pragmatics, Cogo (2009) and finally, pronunciation, which we are focusing on in this paper. The fundamentals of ELF pronunciation can be found generally in the works of Jenkins, who in 2002 proposes a set of 'core' pronunciation features of English taught as an international language, stressing repeatedly in all her works that this EIL pronunciation core should not be identified with the model for imitation in teaching, but rather, should be conceived of as a set of features which teachers should bear in mind when developing the necessary accommodation skills in their students by exposing them to a wide range of varieties of English. Based on these suggestions, core features for individual national ELF varieties have been described, including those for the Croatian 'brand' of ELF (Croglish), as in Josipović (2010).

One of the crucial issues in ELF pronunciation is identity: the division into core and non-core features allows marking one's (national) identity when speaking lingua franca English (cf. e.g. Baxter 1991). For example, Croatian ELF speakers would typically fail to reduce unstressed vowels and speak with syllable-based rhythm, which still would not reduce the international intelligibility of their English. This requires liberal views on pronunciation, specifically, being ready to express one's national identity when speaking English, and accepting others when they express their national identity when speaking English. In other words, a liberal speaker will freely include native-language features in their English and expect others to do the same. A general acceptance of this attitude further suggests acceptance of non-native teaching models. Indeed, all of these were shown to be the characteristics of Croatian learners of English at Croatian universities (Stanojević and Josipović Smojver 2010, 2011). These studies also showed that there was a clear distinction between English majors, potential future teachers of English, who were much less liberal and non-English majors, who were very liberal in their view of English. Thus, overall, their liberality was most significantly related to their choice of major (Stanojević and Josipović Smojver 2011). This is

in line with various qualitative studies, showing that despite the fact that ELF brings about a general feeling of reduced pressure on the learners, they still want to strive for what they believe to be the "correct" target (Jenkins 2006b). Similarly, English teachers in qualitative interviews generally seem opposed to ELF, in the sense that they want to teach "proper" English (Jenkins 2006a, 2009) This kind of ambivalence is referred to by Seidlhofer (2001) as *linguistic schizophrenia*, and further illustrated by Jenkins (2005). It has also been found in Croatian samples of secondary-school learners, as explicitly reported in Kaštelanac (2010) and Poljak (2011).

Overall, this leads to a paradox. University students who are non-English majors are generally liberal with regard to their accents. Some of them will become future speakers of ELF [or "doers", as Firth (2009) refers to them], presumably keeping their liberal views. In contrast, English majors are conservative in relation to their accent. It is conceivable that, as future teachers, they will adhere to their conservative views, and will not be willing to prepare their future students to be speakers of ELF. Rather, they might insist on applying native speaker norms, perpetuating the view of ELF speakers as "deficient communicators" (Firth and Wagner 1997). This suggests that the underlying issue is identity construction, as suggested by a number of studies (e.g. Gatbonton et al. 2005; Jenkins 2005, 2006a). The issue that arises is the following: What is this type of identity construction based on? Is it reasonable to assume that, if it was based on student profile in a university student sample, it will be based on the learner versus speaker status in a more varied sample?

The aim of this paper is to investigate whether the differences behind the attitudes to one's own accent, the accent of one's conversational partners and to teaching models may hinge on belonging to different groups of participants—learners versus 'real' speakers of ELF. We hypothesize that, in addition to some other factors, such as gender, self-assessed proficiency and overall attitude to Croatian, the learner versus speaker status will have a significant impact on the way in which one's own accent, the accent of one's conversational partners, and teaching models are seen. If this proves to be correct, it confirms the claims made by Firth (2009) about the necessity to clearly distinguish between the classroom ELF of learners and the ELF of users belonging to different communities of practice. These two may then be notably different subvarieties of ELF.

The following section gives an overview of research methods. It is followed by the presentation and discussion of the results of the study and a conclusion.

2 Research Methodology

2.1 Sample

In the present paper we will report on a subsample of 623 participants, which consists of 207 employees from a large international company working in Zagreb, a random subsample of 209 secondary school students and a random subsample of

207 university students from throughout Croatia. The two student subsamples were randomly assigned to this study from the entire sample of 2291 university and secondary school pupils, which is being assembled for a larger study. The characteristics of the sample are presented in the text that follows.

2.2 Descriptive Statistics for the Sample

We first look at the participant background. We examined three groups of participants, each accounting for about one third of the entire sample: employees of a large international company (33.2 %), a group of university students (33.5 %), and a group of secondary school pupils (33.2 %).

Of these, 59.1 % of which were female, and 40.6 % were male. Four categories of subjects were looked at: those who have studied English for 9 years and longer, between 5 and 8 years, between 2 and 4 years, and for less than 2 years. With respect to provenance, subjects were divided into three categories: city population, coming from cities with the population of over 50,000 and accounting for 67.7 % of the sample; subjects from towns with the population between 10,000 and 50,000 (21.2 %) and thirdly, village population, which included subjects from places with the population of less than 10,000, who accounted for 11.1 % of the total sample.

The median age of the participants was 25.3. The median age of the employees was 38.3, the median age of the university students was 20.1, and for the secondary school pupils, 17.5.

2.3 Instrument and Procedure

The participants were given an anonymous questionnaire written in Croatian, which consisted of thirty one items. An English translation can be found in Appendix A. In sixteen of these items subjects expressed their attitude by agreement or disagreement with given statements on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree), and in the remaining items they selected one of several options. The first four questions were statements about their attitude to their own regional accent of Croatian, the next seventeen concerned their attitude to English pronunciation, and the remaining ten elicited general information about the subject, such as gender, age, and regional provenance. In this paper we will be reporting only on questions regarding the participants' attitude to their own accent when talking to native and non-native speakers of English (three questions), on two questions regarding their attitudes to their conversational partners, and a single question regarding their attitudes to teaching models. These were tested against the group they belong to (employees, secondary school pupils, university students), self-assessed proficiency, gender, provenance (urban/rural) and attitude to regional recognizability of ideal Croatian. The

questionnaire took about ten minutes to complete. Necessary permissions were obtained from the institutions where the questionnaire was administered. The questionnaire was done on a voluntary basis.

The responses were analysed by means of the Software Package for Social Sciences for Windows (SPSS) 11.0.1. Descriptive statistics, independent samples *t* test, and analysis of variance were used.

3 Results

3.1 Attitudes to One's Own Accent

In this section we will report on whether the participants mind having a foreign accent when talking to native and non-native speakers of English and whether they would want to learn English pronunciation so as to be taken for native speakers.

Overall, most participants (61 %) want to learn English pronunciation so as to be taken for a native speaker, regardless of the time and effort it would take. However, when asked about whether they would mind having a foreign accent when talking to native and non-native speakers of English, most answered they would not mind having a strong or slight accent (76 % when talking to native speakers and 80 % when talking to non-native speakers). This can be interpreted either in terms of Seidlhofer's notion of *linguistic schizophrenia*, referred to above, or simply, as the subjects' realistic assessment of how far one can get in losing one's foreign accent.

We used ANOVA to check for differences between groups on these three questions, specifically differences in participant backgrounds, self-assessed proficiency, gender, regional variety of Croatian and attitudes towards regional recognizability of their Croatian pronunciation.

As for the subjects' background, there are some differences between employees, university students and secondary school pupils with regard to attitudes to one's own accent. Secondary school pupils score lowest on the question of whether they would mind having a strong accent when talking to native or non-native speakers of English (meaning they would mind the least), whereas university students and employees score higher. The difference is not significant with respect to talking to native speakers ($F(2, 613) = 2.49, p = 0.08$), and it is slightly significant on the question related to non-native speakers ($F(2, 614) = 3.52, p = 0.03$). On the latter question, Scheffe's post hoc test showed that secondary school pupils (who scored lowest) were significantly different from the employees (who scored highest), whereas university school students were in the middle. Finally, there was a significant difference between groups on the question whether they would want to speak like native speakers, regardless of the time and effort it would take ($F(2, 618) = 24.19$, $p < 0.001$). Employees score lowest on the question (i.e. wanted least to emulate native speakers), and secondary school pupils score highest. Scheffe's post hoc test

shows that there is a significant difference between employees on the one hand and secondary school and university students on the other.

So university and secondary school pupils do not mind having a strong Croatian accent when talking to native and non-native speakers of English, but score highest on wanting to sound like native speakers. As opposed to that, employees do mind having a strong Croatian accent when talking to native and non-native speakers of English but score lowest on wanting to sound like native speakers.

Self-assessed proficiency is a significant factor in this respect. There are significant differences regarding attitudes to one's own accent when talking to both native ($F(4, 609) = 6.03$, $p < 0.001$) and non-native speakers ($F(4, 610) = 3.47$, $p = 0.008$), with participants who assess themselves as poor scoring lower than participants who assess themselves as excellent. However, Scheffe's post hoc test shows that the difference between groups is not significant. Self-assessed proficiency is not significantly related to wanting to sound like native speakers ($F(4, 614) = 2.26$, $p = 0.06$). In other words, there is a tendency for participants who assess themselves as poorer speakers of English to be less concerned about having a foreign accent. This is something of a chicken-or-egg issue, as it is unclear whether the poor learners feel resigned about their learning capabilities, or whether they are poor learners precisely because they are not ambitious about learning English.

Gender turns out to be a significant factor with regard to one's own accent. There are significant differences between men and women when talking to both native ($t(614) = 3.32$, $p = 0.001$) and non-native speakers ($t(614) = 2.88$, $p = 0.004$), with women minding having a foreign accent more than men (cf. Trudgill 1973, 2000). Also, women score significantly higher on wanting to learn to speak like native speakers regardless of the time and effort it would take ($t(618) = 5.53$, $p < 0.001$). Gender thus seems to cut across the speaker-learner divide. Women want to sound more native-like with both native and non-native speakers, and are willing to take the time and effort to do it.

No significant differences were found between groups of speakers coming from cities, smaller towns or rural communities regarding the attitudes to one's own accent when talking to native ($F(2, 571) = 2.05$, $p = 0.13$) and non-native conversational partners ($F(2, 572) = 0.79$, $p = 0.45$). There was a slightly significant difference between groups on the question of wanting to learn to speak like native speakers regardless of the time and effort it would take ($F(2, 575) = 3.45$, $p = 0.03$), with people living in urban environments scoring lowest, and people from small towns scoring highest, but Scheffe shows no significant differences between the two groups.

Finally, participants who do not believe that ideal Croatian pronunciation should be regionally neutral and regionally unmarked (i.e. they have a liberal attitude to ideal Croatian) do not mind having a foreign accent when talking to native ($F(4, 611) = 7.25$, $p < 0.001$) or non-native speakers of English ($F(4, 611) = 6.69$, $p < 0.001$). The differences are significant. Scheffe's post hoc test shows that there are significant differences between liberal and conservative groups. Importantly, however, there are no significant differences between groups on wanting to learn

English to speak like native speakers ($F(4, 614) = 0.74, p = 0.565$). This means that attitudes to Croatian are related to disfavouring a foreign accent in actual speech, but not in learning English pronunciation.

Overall, the answers to the questions concerning one's own accent suggest that the subjects' liberal attitude in this respect (determined primarily by the gender and to a somewhat smaller degree by the background) implies unlikeliness to be perfectionist about English pronunciation.

Next we checked for any differences regarding the attitudes to the conversational partner's accent. We will focus on two questions—whether the participants prefer talking to non-native speakers of English because they are less worried about their pronunciation being judged, and whether they find it easier to understand the pronunciation of non-native speakers of English. Overall, most participants moderately agree with both statements ($M = 3.11$, $SD = 1.35$; $M = 3.28$, $SD = 1.24$).

Employees score highest on preferring to talk to non-native speakers of English ($M = 3.20$, $SD = 1.39$), and university students score lowest on this question ($M = 2.94$, $SD = 1.38$). We found no significant differences between groups according to the participant background ($F(2,616) = 2.43, p = 0.09$). There was a significant difference between groups on finding it easier to understand the pronunciation of non-native speakers ($F(2, 611) = 23.92$, $p < 0.001$), with Scheffe showing that employees score significantly higher than university students and secondary school pupils. In other words, employees find it easier to understand the pronunciation of non-native speakers, but this does not bring about the preference of talking to non-native speakers because of their pronunciation being judged.

There was a significant difference between groups of speakers on self-assessed proficiency on the two questions—preference to talk to non-native speakers ($F(4, 612) = 12.06$, $p < 0.001$) and easier understanding of non-native speakers ($F(4, 607) = 4.43$, $p = 0.002$). Scheffe's post hoc test shows that speakers who assess themselves as excellent score significantly lower on the first question than all other groups. Scheffe shows no significant differences between groups on the second question. Thus, participants who assess themselves as poorer speakers prefer talking to non-native speakers and find them easier to understand, as opposed to participants who assess themselves better. This coincides with Jenkins' (2000) finding that non-native speakers use particular types of accommodation towards other non-native speakers, which may tend to be different from those used in native speaker—non-native speaker communication. Moreover, there are a variety of studies which found a negative correlation between anxiety and performance on oral tests (e.g. Philips 1992; Mihaljević Djigunović 2007, 121), and studies where talking to native speakers was found to be a significant stressor (e.g. Woodrow 2006). Therefore, it seems that accommodation skills of non-native speakers and language anxiety play a role here.

There were no significant differences between men and women on preferring to talk to non-native speakers ($t(616) = 1.87$, $p = 0.06$). Women find it somewhat easier to understand non-native speakers, and the difference is slightly significant ($t(612) = -2.30$, $p = 0.02$).

Finally, there were no significant differences between speakers with a different provenance on the first question ($F(2,574) = 1.44, p = 0.24$), and the results for the second question were on the verge of significance ($F(2, 569) = 3.12, p = 0.05$).

In short, it is self-assessed proficiency, rather than gender of provenance that significantly determines preferences with respect to the nativeness of the conversational partner, which only strongly suggests that ELF is more internationally intelligible than native English, as indicated by the case studies of miscommunication referred to by Canagarajah (2007). Although Canagarajah accounts for this simply as the failure of native speakers to negotiate, even a quick glance at the list of core pronunciation features of ELF identified by Jenkins (2002) clearly shows that ELF forms are regularly unmarked in the Jakobsonian sense, that is, easier to pronounce and perceive.

3.2 Attitudes to Teachers

Finally, we asked about the participants' preference for native or non-native teachers of English to teach them pronunciation. Overall, most participants slightly prefer native speakers as their pronunciation teachers ($M = 3.37$, $SD = 1.20$).

There was a significant difference between participants of different backgrounds ($F(2, 611) = 10.32, p < 0.001$). Scheffe's post hoc test showed that there was a significant difference between employees on the one hand (who scored significantly higher on this question) and university and secondary school pupils on the other (who scored significantly lower).

Self-assessed proficiency proved not to be significant in this respect ($F(4, 607) = 1.43, p = 0.22$). Gender was not significant either ($t(612) = -1.04$, $p = 0.30$), and neither was the participants' provenance ($F(2, 569) = 2.04$, $p = 0.13$).

Overall, the only significant predictor of the attitude to the native versus non-native status of the teacher is the subjects' background. Interestingly, the learner versus user distinction turns out to be decisive.

4 Discussion and Conclusion

The present findings result from different identity constructs. The obtained results tie in with the findings of Stanojević and Josipović Smojver (2011), which show that the students' major subject is the most significant predictor of the attitude to nativeness versus non-nativeness in English pronunciation, even though other factors are at play, such as gender and self-assessed proficiency. In the present research we find an analogous situation. Once again, variables such as gender and self-assessed proficiency turn out to influence the subjects' attitude to English pronunciation.

However, an additional finding presented and discussed in this paper is that it is primarily the learner versus user role that determines one's striving for native-like pronunciation or settling for ELF without bothering much about one's foreign accent. This corresponds to the dichotomy dealt with by Firth and Wagner (1997) and Canagarajah (2007). Along these lines, our subjects who qualify as 'learners' (secondary-school pupils and university students) construct their speaker identity differently than 'doers' i.e., ELF users. Interestingly enough, when asked questions related to the other, users' role, learners express extreme attitudes: they reportedly do not care how they pronounce English in conversation outside classroom, i.e., they are not bothered by their Croatian accent. This ties in with the fact that they prefer, that is, find it easier to speak to non-natives, even though they are not the ones who mostly use English in conversation in natural situations.

As opposed to what we have identified as learners, the employees are real speakers, or 'users', who then construct their role accordingly. They are bothered by a strong accent when speaking English, no matter whether the conversational partners are native speakers or not. In other words, they are embarrassed about what they feel as a 'bad accent'. This is hardly surprising, considering the other, independent results, according to which Inner-Circle accents are good and desirable. They also prefer to talk to non-natives, which they find to be not only less embarrassing, but notably easier to understand. As they are not involved in learning English in formal instructional settings they do not construct the opposite role, i.e., that of learners, "realistically", but tend to be extreme in their views of how learning should proceed. Thus, they do not want to waste time and effort on learning 'nice', native-like English pronunciation. Nevertheless, they believe they would learn to pronounce English better from a native speaker. As Firth (2009, 127) nicely puts it, they "disavow any intimations of 'learner' status" and believe that they will learn English more successfully if they do it through communication with natives outside the classroom.

In the present results there are some indications that both subcategories of subjects—pupils and university students, i.e., learners on the one hand and employees on the other, are willing to accept the reality of ELF. Still, these indications vary with the roles the subjects assume. As learners, they are not bothered by their foreign accent, although they assess native, Inner-Circle varieties as superior. As users outside the classroom, they are not willing to improve upon their pronunciation, but they still 'respect' Inner-Circle pronunciation and prefer it in teaching models. In addition, they understand non-native speakers better and prefer them as conversational partners.

Theoretically speaking, different identity constructions with regard to ELF warrant a stratified view of ELF: i.e. ELF is regarded differently by people with different roles and at different stages in their lives. Speakers of ELF do not form an "imagined community" (Prodromou 2008, 89), because ELF is not characterized by a stable "horizontal comradeship" (Anderson 1991, 7). Its stratification is visible not only in the number of ELF variants (Croglish, Spanglish, etc.) but also in ELF users' attitudes to it. This may be the reason why ELF must be mediated to learners, teachers and scholars: it is not based on a unified national machinery,

whose "comradeship", among other things, results in clear attitudes about what constitutes "good" and "bad"/"proper" and "improper" linguistic usage. It is not monolithic (Jenkins 2006a, 161) and stable: instead its stability is constructed between speakers in particular communicative situations (where they may regard themselves as speakers rather than learners; cf. e.g. Erling and Bartlett 2006; Firth 2009). As scholars and theoreticians we should constantly keep this in mind.

As teachers of English, being aware of the dynamism and stratification of ELF, we should expose our students to a wide range of English varieties, including ELF (though certainly not imposing them as models for imitation) and teach them to be flexible about using English in real-life situations. This implies developing all kinds of accommodation strategies, such as repetition, code-switching or recognizing, or (if necessary) resorting to somewhat simplified ELF rather than native-like forms, drawing on the speakers' multilingual and multicultural repertoire (cf. Cogo 2009). ELF is a reality, but an elusive one, and although our students (who define themselves as learners) may not approve of ELF in their current learner role, they may require it later when they start defining themselves as speakers.

Finally, if ELF turns out to become reality in Croatian, in addition to being aware of it and teaching our learners all those strategies necessary for an effective use of it, we must achieve some other goals as well. In the first place, we must extensively describe and define the Croatian core ELF. Besides, we must try to change attitudes to it, making it clear that it does not constitute a monolithic imitation model but a dynamic variety emerging in communication. Introducing university ELF courses for future English teachers would also be a way to prepare new generations of teachers and ultimately, future ELF users for the effective use of English as a global language.

Acknowledgments We wish to thank Moira Kostić Bobanović, Snježana Kereković, Višnja Kabalin Borenić, Mario Brdar, Zrinka Jelaska, Kristina Cergol, Jelena Parizoska, Ivanka Rajh, Ana Poljak, Anamarija Kaštelanac, Žarko Magdić, Maja Bradić, Marina Mioč, Katarina Seljan, and Marijana Pejaković for their help in administering the questionnaire.

Appendix A: Questionnaire About the Pronunciation of Croatian and English

Please fill in this anonymous questionnaire, which examines the attitudes concerning the pronunciation of Croatian and English. The questionnaire consists of two parts: four statements about Croatian and seven statements about English. Please circle the answer which best corresponds to your opinion. When you are asked to assess a statement on a 5-point scale, please use the following values:

Strongly disagree	Partially disagree	Neither agree nor disagree	Partially agree	Strongly agree
1	2	3	4	5

Part 1: attitudes to Croatian pronunciation

1. I would like my pronunciation of Croatian to be regionally recognizable (e.g. as coming from the Zagreb area, as Kajkavian, Slavonian, Dalmatian…).

1 2 3 4 5

1. When public figures (e.g. singers, politicians, actors, etc.) speak on the radio and TV:

a. I find it acceptable when I can tell by their Croatian pronunciation (accent) where they are from, whatever accent they have

b. I find it acceptable when I can tell by their Croatian pronunciation (accent) where they are from, but only for some Croatian accents, and not for others

c. When public figures speak, any regional pronunciation (accent) bothers me.

2. When news readers speak on the radio and TV:

a. I find it acceptable when I can tell by their Croatian pronunciation (accent) where they are from, whatever accent they have

b. I find it acceptable when I can tell by their Croatian pronunciation (accent) where they are from, but only for some Croatian accents, and not for others

c. When news readers speak, any regional pronunciation (accent) bothers me

3. I believe that ideal Croatian pronunciation should be regionally neutral and regionally unmarked.

1 2 3 4 5

Part 2: attitudes to English pronunciation

4. When I speak English, I believe that it is important:

a.	that I am fluent	**1**	**2**	**3**	**4**	**5**
b.	that my pronunciation is correct	**1**	**2**	**3**	**4**	**5**
c.	that my grammar is correct	**1**	**2**	**3**	**4**	**5**

5. When I speak English with native speakers (e.g. the English or Americans):

a. I do not mind having a strong Croatian accent

b. I do not mind having a slight Croatian accent

c. I do not want to speak with a foreign accent.

6. When I speak English with non-native speakers (e.g. the French, Italians, Germans):

a. I do not mind having a strong Croatian accent

b. I do not mind having a slight Croatian accent

c. I do not want to speak with a foreign accent.

7. I prefer speaking English with other non-native speakers of English (e.g. Germans, Italians, the French) than with native speakers of English (e.g. the English, Americans) because when I speak with non-native speakers, I am not worried about them judging my pronunciation.

1 2 3 4 5

8. If I could perfect my English pronunciation so as to pass for a native speaker, I would do it regardless of the time and effort it would take.

YES NO

9. A. If you answered YES to question 9: Why would you perfect your English pronunciation?

a. because it significantly improves the general impression of me and my knowledge of English

1 2 3 4 5

b. because I like to impress my conversational partners

1 2 3 4 5

c. because it is important for my job

1 2 3 4 5

d. because I am generally a perfectionist, and that includes my pronunciation **1 2 3 4 5**

9. B. If you answered NO to question 9: Why would you not perfect your English pronunciation?

a. because native speakers dislike foreigners who try too hard to sound like them **1 2 3 4 5**

b. because I want to preserve my national identity when I speak a foreign language **1 2 3 4 5**

c. because learning the proper pronunciation is a waste of time which can be better spent on other goals such as learning grammar and vocabulary

1 2 3 4 5

d. because people will think that I am putting on airs

1 2 3 4 5

9. I find it easier to understand the pronunciation of non-native speakers of English (e.g. Germans, Italians, the French) if they are fluent and if their grammar is correct than the pronunciation of native speakers of English.

1 2 3 4 5

10. I believe I will learn to pronounce English better if I am taught by an English teacher who is a native speaker than by an English teacher who is Croatian, and who speaks English as a foreign language.

1 2 3 4 5

Participant data

11. Gender: F M
12. Age:
13. Schooling: secondary school junior college university
14. Where did you spend most of your life (which village, town)? _____
15. How long have you been learning English?

less than 2 years 2-4 years 5-8 years 9 years and longer

16. Assess your knowledge of English:

a) very poor b) poor c) good d) very good e) excellent

17. Assess your pronunciation of English:

a) very poor b) poor c) good d) very good e) excellent

18. How often do you speak English?

a) hardly ever b) several times a year c) once a month d) every week

19. Have you ever spent more than 6 months in an English speaking country?

YES NO

If so, where? _____

References

Anderson, Benedict. 1991. *Imagined communities: reflections on the origin and spread of nationalism*. Revised and extended edition. London: Verso.

Baxter, James. 1991. How should I speak English? American-ly, Japanese-ly or internationally. In *Teaching English pronunciation: a book of readings*, ed. Adam Brown, 53–71. London: Routledge.

Canagarajah, Suresh. 2007. Lingua franca English, multilingual communities, and language acquisition. *The Modern English Journal* 91:923–939.

Cogo, Alessia. 2009. Accommodating difference in ELF conversations: A study of pragmatic strategies. In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 254–273. Newcastle upon Tyne: Cambridge Scholars Publishing.

Crystal, David. 2004. *The language revolution*. Cambridge: Polity Press.

Dewey, Martin. 2007. English as a lingua franca and globalization: An interconnected perspective. *International Journal of Applied Linguistics* 17:332–354.

Dewey, Martin. 2009. English as a lingua franca: Heightened variability and theoretical implications. In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 60–83. Newcastle upon Tyne: Cambridge Scholars Publishing.

Ehrenreich, Susanne. 2009. English as a lingua franca in multinational corporations – exploring business communities of practice. In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 126–151. Newcastle upon Tyne: Cambridge Scholars Publishing.

Erling, Elizabeth J, and Tom Bartlett. 2006. Making English their own: The use of ELF among students of English at the Free University of Berlin. *Nordic Journal of English Studies* 5: 9–40.

Firth, Alan. 2009. Doing not being a foreign language learner: English as a lingua franca in the workplace and (some) implications for SLA. *IRAL—International Review of Applied Linguistics in Language Teaching* 47:127–156.

Firth, Alan, and Johannes Wagner. 1997. On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal* 81:285–300.

Gatbonton, Elizabeth, Pavel Trofimovich, and Michael Magid. 2005. Learners' ethnic group affiliation and L2 pronunciation accuracy: A sociolinguistic investigation. *TESOL Quarterly* 39:489–511.

Görlach, Manfred. 2002. *Still more Englishes*. Amsterdam: John Benjamins.

Jenkins, Jennifer. 2000. *The phonology of English as an international language: new models, new norms, new goals*. Oxford: Oxford University Press.

Jenkins, Jennifer. 2002. A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics* 23:83–103.

Jenkins, Jennifer. 2005. Implementing an international approach to English pronunciation: The role of teacher attitudes and identity. *TESOL Quarterly* 39:535–543.

Jenkins, Jennifer. 2006a. Current perspectives on teaching world Englishes and English as a lingua franca. *TESOL Quarterly* 41:157–181.

Jenkins, Jennifer. 2006b. English pronunciation and second language speaker identity. In *The sociolinguistics of identity*, ed. Tope Omoniyi and Goodith White, 75–91. London, New York: Continuum.

Jenkins, Jennifer. 2007. *English as a lingua franca: Attitude and identity*. Oxford: Oxford University Press.

Jenkins, Jennifer. 2009. English as a lingua franca: interpretations and attitudes. *World Englishes* 28: 200–207.

Josipović Smojver, Višnja. 2010. Foreign accent and levels of analysis: interference between English and Croatian. In *Issues in accents of English 2: variability and norm*, ed. Ewa Waniek-Klimczak, 23–35. Newcastle upon Tyne: Cambridge Scholars Publishing.

Kaštelanac, Anamarija. 2010. The attitude of secondary-school pupils from Zadar to English pronunciation. (ms.) Unpublished graduation thesis. University of Zagreb.

Mauranen, Anna, and Elina Ranta. eds. 2009. *English as a lingua franca: Studies and findings*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Mihaljević Djigunović, Jelena. 2007. Croatian EFL learners' affective profile, aspirations and attitudes to English classes. *Metodika* 8: 115–126.

Phillips, Elaine M. 1992. The effects of language anxiety on students' oral test performance and attitudes. *The Modern Language Journal* 76: 14–26.

Pitzl, Marie-Luise. 2009. 'We should not wake up any dogs': Idiom and metaphor in ELF. In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 298–322. Newcastle upon Tyne: Cambridge Scholars Publishing.

Poljak, Ana. 2011. Attitudes of secondary-school EFL learners towards the pronunciation of English. (ms.) Unpublished graduation thesis. University of Zagreb.

Prodromou, Luke. 2008. *English as a lingua franca: a corpus-based analysis*. London, New York: Continuum.

Pullin Stark, Patricia. 2009. No joke—this is serious! Power, solidarity and humour in Business English as a Lingua Franca (BELF). In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 152–177. Newcastle upon Tyne: Cambridge Scholars Publishing.

Ranta, Elina. 2009. Syntactic features in spoken ELF—learner language or spoken grammar? In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 84–106. Newcastle upon Tyne: Cambridge Scholars Publishing.

Seidlhofer, Barbara. 2001. Closing a conceptual gap: the case for a description of English as a lingua franca. *International Journal of Applied Linguistics* 11:133–158.

Seidlhofer, Barbara. 2009. ELF findings: form and function. In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 37–59. Newcastle upon Tyne: Cambridge Scholars Publishing.

Smit, Ute. 2009. Emic evaluations and interactive processes in a classroom community of practice. In *English as a Lingua Franca: Studies and Findings*, ed. Anna Mauranen and Elina Ranta, 200–224. Newcastle upon Tyne: Cambridge Scholars Publishing.

Stanojević, Mateusz-Milan, and Višnja Josipović Smojver. 2010. Euro-English and Croatian national identity. Conference paper presented at the New Challenges for Multilingualism in Europe, Dubrovnik, Croatia, April 11–15 2010.

Stanojević, Mateusz-Milan, and Višnja Josipović Smojver. 2011. Euro-English and Croatian national identity: Are Croatian university students ready for English as a lingua franca? *Suvremena lingvistika* 37: 105–130.

Trudgill, Peter. 1973. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, Peter. 2000. *Sociolinguistics: An introduction to language and society*. Penguin.

Woodrow, Lindy. 2006. Anxiety and Speaking English as a Second Language. *RELC Journal* 37: 308–328.

Intonation in the Perception of Brummie

Kamil Malarski

Abstract The Birmingham accent, also known as Brummie, enjoys a very bad reputation in Great Britain. It was suggested that its intonation is responsible for the stigma in the first place (How to speak Brummie, http://www.bbc.co.uk/dna/h2g2/A496352, 2001). Since Birmingham intonation differs from standard British intonation, and because intonation is indeed pragmatically meaningful for the perception of speech, two experiments at Adam Mickiewicz University were carried out to verify the hypothesis. The first one was to answer if Brummie is indeed disfavoured when compared to other dialects. The subjects listened to three Brummie speakers and three speakers of different accents of English, and rated the perceived attractiveness, friendliness and intelligence of the recordings on a 5-point Likert scale. The accent, as expected, was deemed the least attractive and intelligent. The second, core part of the experiment investigated to what extent intonation is responsible for this bad perception of Brummie speech. This time, the subjects were to listen to two versions of the speech sample of a given accent. The first version contained only intonation (the speech was unintelligible), and the second one included only segmentals (intonation was removed from the signal). The modifications were made using the methodology of Van Bezooijen and Gooskens (J Lang Soc Psychol 18(1):31–48, 1999) in PRAAT. The results of the two versions of each recording were then compared. On the whole, it turned out that Brummie intonation is indeed seen more negatively than RP intonation.

K. Malarski (✉)
Adam Mickiewicz University, Poznań, Poland
e-mail: kamil_go@yahoo.com

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_15,

1 Introduction

Brummie is the accent spoken in the city of Birmingham in the area of the West Midlands. Despite its being widely discussed in the media, Brummie has not received too much attention from linguists. This was noted by Foulkes and Docherty (1999) and several years later by Clark (2008). There are a few sources (Wells 1982; Clark 2008) describing its segmental phonology but the information about its suprasegmental features is really scarce. The following paper is an attempt to fill this gap and to understand popular opinions on this accent. This article is largely based on Malarski (2010).

2 Brummie's Stigma

The West Midlands accent is possibly the most stigmatized variety in Great Britain. It scores worst in various accent evaluation studies that considers criteria like intelligence, prestige or attractiveness (Giles 1970; Hiraga 2005; Coupland and Bishop 2007). Among other things, it has been described by people as "lazy", "ugly" and "uneducated" (Thorne 2005). This has serious social implications for the Brummie users. As Dixon et al. (2003) show, subjects are much more likely to be thought guilty when accused of a crime if they speak in Brummie rather than in Standard British English. Giles et al. (1975), too, show that people usually see the Birmingham accent speakers as less trustworthy than the RP speakers.

Of course, there is no one simple answer to why certain varieties of English are perceived as worse than others. The bias against the Birmingham accent could be because it has traditionally been an industrial city with a large proportion of inhabitants coming from lower socioeconomic classes. Apart from the obvious connection with the people who speak them, accents are also evaluated on the basis of their phonetic traits (Van Bezooijen and Gooskens 1999, 47). It was suggested in an online article (http://www.bbc.co.uk/dna/h2g2/A496352) that Brummie is seen negatively entirely because of its intonation. The article was written by a layperson rather than a linguist but strongly inspired the following study.

3 Intonation in Brummie

Linguists have emphasised that intonation is pragmatically and semantically meaningful for the listeners (Grabe et al. 2003, 379) and conveys emotion (Hlebec 2008). It has even been suggested that prosody, which includes intonation, is more telling than segments in evaluating the speech of others (Anderson-Hsieh et al. 1992).

Brummie intonation differs from the Standard British English intonation, primarily in that Brummie and other urban British varieties like Scouse, Geordie, Belfast and Glasgow use rising tones for statements (Wells 1982, 91; Cruttenden 1994, 138–139). These rises, in Brummie, are mostly realized by "rise-plateaus" and "rise-plateau-slumps" (Cruttenden 1994, 139). Cruttenden (1994, 139) explains his own term "rise-plateau" as "the jump-up on the unaccented syllable following the nucleus and the maintenance of this level on succeeding unaccented syllables". The "rise-plateau-slump" is similar with the exception that the last two syllables in an intonational unit can drop in pitch.

Ladd (1996, 125) interprets Birmingham declaratives as demonstrating a stylised low rise intonation and asserts that they are the same tones as rise-plateau-slumps. Brummie questions, on the other hand, are interpreted as rise-falls (Ladd 1996, 125). These intonational contours, that is low rises and rise-falls, appear in standard British intonation. There, however, they convey negative attitudes. In RP, low rises communicate "resentfulness", "deprecation", "reproving criticism" (O'Connor and Arnold 1967, 169), "non-finality" and "suggestion" (Hirst 1999, 63). Rise-falls, in questions, convey "challenging", "antagonistic" and "disclaiming responsibility" attitudes (O'Connor and Arnold 1967, 147). Thus, it seems that there are reasonable grounds to claim that Brummie rising tones can elicit unintended negative interpretations.

4 Experiment 1

The experiment comprises of two parts. In the first experiment (Experiment 1), the subjects were asked to evaluate three Brummie speakers and three speakers of different varieties of English according to three criteria. Its aim was to discover whether the Birmingham accent is as disfavoured by Polish students as it is among British people. The first part was meant as an introduction to the second and more important part of the experiment.

4.1 Data

Six speech samples appeared in Experiment 1. There were three Brummie speakers (all males), one RP speaker (a male), one rural West Yorkshire speaker (a female) and one Liverpool speaker (a female). Two Brummie recordings came from the British Library Archival Sound Recordings available online (1998). One Brummie speaker was recorded from the Internet. The remaining samples came from the International Dialects of English Archive website (1997). All six voices were recorded using PRAAT 5.1.18. For better loudness and quality, they were normalised and dynamically compressed in Audacity 1.2.6. The samples lasted from 9 to 18 s.

4.2 Procedure

The experiment was in the form an online questionnaire carried out through the Moodle e-learning platform at the departmental website of the School of English at Adam Mickiewicz University in Poznań. The participants were to evaluate the recordings on the 5-point Likert scale. The criteria were niceness, intelligence and friendliness, features which are usually low-scoring for Brummie. On the scale, 1 meant ugly/unintelligent/unfriendly whereas 5 meant nice/intelligent/friendly. 3 was neutral. The listeners were also given the chance to comment in their own words on how a speaker sounded. Spoken passages which could tell something about the origins of the speakers, or their occupation, or in any way influence the judgments were eliminated. The questionnaire can be found in Appendix 1.

4.3 Subjects

Thirty-one Polish students of English at the School of English at Adam Mickiewicz University in Poznań, 26 women and five men, took part in the experiment. They were judged to be very fluent in English and ranged in age from 21 to 26. Twelve were familiar with the American pronunciation model, and 19 had learnt British pronunciation. Although they had had formal training in phonetics and phonology and linguistics, they were not expected to identify the dialects, which were presented in this study, apart from RP (Standard Southern British) (see, e.g., Weckwerth et al. 2005 for accent recognition among Polish students of English).

4.4 Hypothesis

It was argued in the preceding sections that certain intonational features of Brummie have the potential to influence people's negative reactions towards the accent. Therefore, Birmingham English is expected to score low in this accent evaluation study.

5 Results

The complete results are presented in Table 1 of Appendix 2. The scores were rounded to the nearest 0.1. The accents were classified according to the overall result, which was the sum of the scores from each of the three criteria. Additionally, at the bottom of the table, there appears the mean score for all three Brummie speakers.

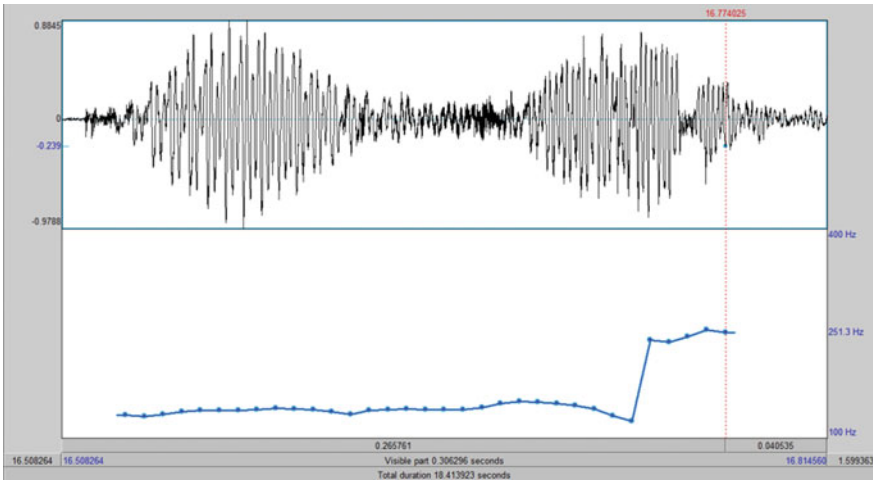


Fig. 1 The sentence-final word "brother" pronounced by Brummie 1 speaker

The Birmingham accent, in average, was judged to be less desirable than other accents which confirms the hypothesis for this part of the experiment. Two Birmingham speakers were evaluated as least acceptable. However, one Brummie user scored a little better and came before the Liverpool speaker. Birmingham English was judged to be an unintelligent and rather ugly accent. Intelligence was ranked near the bottom, and it is the only mean Brummie score that is below 3.

There is one especially cogent finding for the role of intonation in perceiving the Birmingham accent: one subject commented, "the last segment in *brother* scared me". The word "brother" was the last word in the utterance and it was said on a rising intonation. On the spectrogram, it resembles a rise-plateau (see Fig. 1). What the listener meant was probably not the last *segment* (which was a schwa) but the last syllable, which had a characteristic rising intonation.

6 Experiment 2

The second part of the experiment, inspired by the work of Van Bezooijen and Gooskens (1999), forms the core of this paper. It was aimed at determining exactly to what extent Brummie intonation is judged negatively.

6.1 Data

In this part of the experiment, there were twelve recordings to be listened to by the participants. These were six speech samples, each played in two versions. The first

version included only intonation, the second version the segmental information with a flat intonation. To attain an intonational version of a speech sample a lowpass filter at 350 Hz was used. Thus, the speech was unintelligible, and only intonation was heard. In the segmental version, pitch was flattened and normalised at 109 Hz, i.e., an average pitch level for all speakers. All modifications were performed in PRAAT using the methodology of Van Bezooijen and Gooskens (1999). The recordings, just like in the first part of the experiment, were compressed and normalised for better quality. They lasted, with one exception, from 10 to 30 s, which was considered long enough to carry a recognisable intonation pattern. Even in the shortest speech sample, which lasted less than 2 s, the two sentences were included. The samples came from the same speakers which were featured in Experiment 1.

6.2 Procedure

The speakers appeared in a different order than they did in Experiment 1. Also, the samples were different. Two versions of each speech sample were played one by one, the intonational version first and the flat-intonation version second. The listeners were supposed to rate each of them on exactly the same Likert scale as in the first part of the experiment.

6.3 Subjects

The same subjects, who participated in Experiment 1, took part in Experiment 2.

6.4 Hypothesis

My hypotheses were that the Birmingham intonation would be less favoured than RP intonation and that the intonational versions of Brummie speech would be less favoured than the ones with flat intonation.

7 Results

The complete results are presented in Tables 2 and 3 in Appendix 2. The scores, again, were rounded to one decimal place.

The first hypothesis is easily supported. RP intonation was seen as friendlier, more intelligent and nicer than Birmingham intonation. All three Brummie

speakers came at the bottom of the table which means that their intonation was considered less attractive than the intonation of other varieties. To either support or reject the second hypothesis, however, the results for each Birmingham speaker have to be investigated separately.

Results for Speaker Brummie 1 suggest that the utterances with normal intonation were more acceptable than those with flat intonation. Here, the hypothesis that the rising tones at the end of declarative sentences evoke unpleasant feelings in the listeners is not applicable, for the speaker used a rise sentence-finally. The pattern (see Fig. 2) resembles a rise-plateau where the relatively high pitch of the rise is maintained until the end of the second sentence.

For the second Birmingham speaker (Brummie 2), on the contrary, the hypothesis is confirmed as his utterances with normal intonation was perceived as uglier and less friendly than his utterances with flat intonation. This speaker, however, produced a falling tone sentence-finally. He did produce some rising tones in statements but in non-final positions which is very common in standard speech. Therefore, although on the surface it appears that the hypothesis was confirmed, this speaker did not use contours common in Brummie.

The third Brummie user (Brummie 3) scored exactly the same overall for both types of utterance. His rank for utterances with normal intonation was the lowest of all speakers (see Table 2). These negative judgments were not caused by the use of a rising tone sentence-finally because the speaker used a falling tone at the end of the sentence. Here, the second hypothesis is rejected, as well.

8 Conclusion

It has been suggested by lay observers that the low status of Birmingham English may be due to the intonation used in this accent. Linguistic explanations also point to intonation as very important in understanding and evaluating the speech of others. Birmingham intonation differs from standard Southern British intonation. Experiment 1 shows that rising tones in final positions as used in Brummie can be unpleasant for the subjects. In Experiment 2, Birmingham intonation was arguably viewed by the listeners as uglier, less friendly and less intelligent than the standard intonation. However, results were somewhat ambiguous in this respect.

As noted above, literature on the Birmingham accent is scant. This study is a small contribution, but there are numerous possible directions for further research. The Brummie pitch range, which can have a great influence on how speech is perceived (see, e.g., Grabe et al. 2003, 379), is an obvious example.

Acknowledgments I would like to thank dr Jarosław Weckwerth for very helpful comments on the earlier draft of this article.

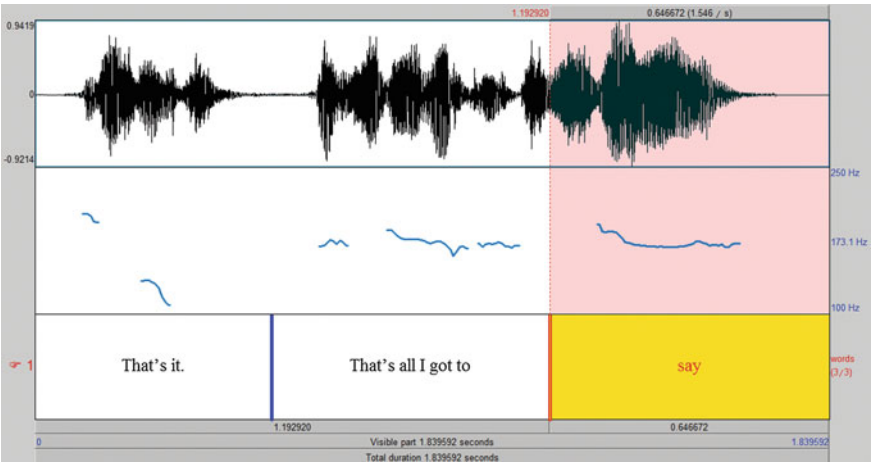


Fig. 2 The rising intonation by Brummie 1 speaker

Appendix 1

Questionnaire on accents of English

Please take 10 min to fill in this questionnaire

Please answer these questions about yourself first

1. Are you male or female?
2. How old are you? [optional]
3. A student of which year are you currently?
4. How long have you been learning English?
5. Which pronunciation model have you been taught?
6. Have you been to an English speaking country? How many times? How long did you stay there? Where exactly have you been?

Listen to the following speech samples and rate them on the scales provided. Preferably, listen to these through headphones. You can also leave some comments.

7. Listen to the recording.

   Does the speaker sound
   1-Ugly, 5-Nice
   1-Unfriendly, 5-Friendly
   1-Unintelligent, 5-Intelligent

8. Comment on how the speaker sounds if you like.

   [The same instructions for the remaining five recordings].

Now, listen to these recordings. They have been modified in some ways. Rate them on the same scales. Preferably, listen to these through headphones. You can also leave some comments.

19. Listen to the recording.

    Does the speaker sound
    1-Ugly, 5-Nice
    1-Unfriendly, 5-Friendly
    1-Unintelligent, 5-Intelligent

20. Comment on how the speaker sounds, if you like.

    [The same instructions for the remaining eleven recordings].

43. If you have any comments on the questionnaire write them down.

    Thank you very much for devoting your time and filling in this questionnaire.

Appendix 2

Table 1 Students' reactions to Brummie and other accents according to the 5-point Likert scale

Accent	Niceness	Friendliness	Intelligence	Overall
RP	4.4	4.2	4.0	12.6
Rural West Yorkshire	3.8	3.9	3.5	11.2
Brummie 2	3.6	3.8	3.3	10.7
Liverpool	3.5	3.1	3.4	10.0
Brummie 1	3.3	3.5	3.0	9.8
Brummie 3	2.3	3.0	2.5	7.8
Mean Brummie	3.1	3.4	2.9	9.4

Table 2 Students' attitudes towards intonation of the six varieties of British English

Accent	Niceness	Friendliness	Intelligence	Overall
Rural West Yorkshire intonation	2.5	3.1	2.8	8.4
RP intonation	2.5	2.8	2.9	8.2
Liverpool intonation	2.5	2.6	2.7	7.8
Brummie 2 intonation	2.3	2.6	2.8	7.7
Brummie 1 intonation	2.2	2.3	2.5	7.0
Brummie 3 intonation	2.1	2.3	2.5	6.9

Table 3 Students' attitudes towards segments of the six varieties of British English

Accent	Niceness	Friendliness	Intelligence	Overall
RP segments	2.9	2.9	3.1	8.9
Brummie 2 segments	2.7	2.9	2.6	8.2
Brummie 3 segments	2.1	2.5	2.3	6.9
Liverpool segments	2.0	2.4	2.3	6.7
West Yorkshire segments	1.7	2.2	2.5	6.4
Brummie 1 segments	1.9	1.9	2.5	6.3

References

Anderson-Hsieh, Janet, Ruth Johnson and Kenneth Koehler. 1992. The relationship between native speaker judgements of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning* 42: 529–555.

British Library Archival Sound Recordings. 1998. (http://sounds.bl.uk/SearchResults.aspx?query=Birmingham&category=Accents-and-dialects&publicdomainonly=false) (date of access: 5 Feb. 2010).

Clark, Urszula. 2008. The English West Midlands: phonology. In *Varieties of English 1: The British Isles*, eds. Bernd Kortmann and Clive Upton, 145–177. Berlin: Mouton de Gruyter.

Coupland, Nikolas and Hywel Bishop. 2007. Ideologised values for British accents. *Journal of Sociolinguistics* 11: 74–93.

Cruttenden, Alan. 1994. *Intonation*. Cambridge: Cambridge University Press.

Dixon, John A., Berenice Mahoney and Roger Cocks. Accents of guilt? Effects of regional accent, race, and crime type on attributions of guilt. *Journal of Language and Social Psychology* 21: 162–168.

Foulkes, Paul and Gerard J. Docherty. 1999. *Urban voices: Accent studies in the British Isles*. London: Arnold.

Giles, Howard. 1970. Evaluative reactions to accents. *Educational Review* 22: 211–227.

Giles, Howard, Susan Baker and Guy Fielding. 1975. Communication length as a behavioural index of accent prejudice. *Linguistics* 166: 73–81.

Grabe, Esther, Burton S. Rosner, José E. García-Albea and Xiaolin Zhou. 2003. Perception of English intonation by English, Spanish, and Chinese listeners. Language and Speech 46: 375–401.

Hiraga, Yuko. 2005. British attitudes towards six varieties of English in the USA and Britain. *World Englishes* 24: 289–308.

Hirst, Daniel. 1999. Intonation in British English. In *Intonation Systems. A Survey of Twenty Languages*, eds. Daniel Hirst and Albert Di Cristo, 56–77. Cambridge: Cambridge University Press.

Hlebec, Boris. 2008. General attitudinal meanings in RP intonation. *Studia Anglica Posnaniensia* 44: 275–295.

International Dialects of English Archive. 1997.
(http://web.ku.edu/~idea/europe/england/england.htm) (date of access: 5 Feb. 2010).

How to speak Brummie. 2001. (http://www.bbc.co.uk/dna/h2g2/A496352) (date of access: 25 Nov. 2009).

Ladd, Robert D. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

Malarski, Kamil. 2010. The Role of intonation in the perception of Brummie. Unpublished B.A. paper. Adam Mickiewicz University.

O'Connor, Joseph D. and Gordon F. Arnold. 1967. *Intonation of Colloquial English*. London: Longmans.

Thorne, Steve. 2005. Accent pride and prejudice: Are speakers of stigmatized variants really less loyal? *Journal of Quantitative Linguistics* 12: 151–166.

Van Bezooijen, Renée and Charlotte Gooskens. 1999. Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology* 18, 1: 31–48.

Weckwerth, Jarosław, Katarzyna Janicka and Małgorzata Kul. 2005. Polish students' attitudes to native English accents as models for ESL pronunciation. In *Pronunciation models: A changing scene?*, eds. K. Dziubalska-Kołaczyk and J. Przedlacka, 251–292. Frankfurt/Main: Peter Lang Verlag.

Wells, John C. 1982. *Accents of English.* New York: Cambridge University Press.

Sociophonetic Transfer in Polish Learners of English: The Case of Nasals

Anna Gralińska-Brawata

Abstract Transferring Polish nasal vowels in contexts where English has a combination of vowel + nasal consonant followed by a fricative is among a considerable number of pronunciation problems that Polish learners of English encounter in their learning process. Seemingly a less disturbing and often unnoticed feature of Polglish, it is believed to be a very frequent phenomenon (Sobkowiak, 2004, 204. *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie), which seems to result from the systemic differences between the two languages. The present study reports on the experiment conducted among 50 learners of English reading texts in English and Polish: *The North Wind and the Sun* and a Polish text of equal length and similar rhythmic structure. Two words, "considered" and "mądrych" were extracted from the English and Polish recorded material and examined using both spectrographic picture and auditory analysis. The results confirm that transferring a Polish nasal vowel /õ/ in contexts where English has /ɒn/ or /ən/ followed by a fricative is a common error among Polish learners of English. What prevails in the pronunciation of the word "mądrych" is /ɔn/ instead of /õ/, which validates the general tendency of Polish speakers to pronounce a nasal vowel before fricatives and a combination of vowel + nasal consonant before plosives and affricates. Statistical tests employed in the study revealed that there is a significant systematic difference between English and Polish in terms of nasal transfer. From the pedagogical point of view, it can be worthwhile to raise awareness among practical phonetics teachers as regards variability in Polish while developing their students' phonological system of English.

A. Gralińska-Brawata (✉)
University of Łódź, Łódź, Poland
e-mail: agrabra@uni.lodz.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_16,

1 Introduction

Polish learners of English often feel downcast by the pronunciation difficulties that may abound in the course of their learning process. These problems, both on segmental and suprasegmental levels, may be rated as more or less relevant in terms of intelligibility or attaining native-like pronunciation by means of various evaluation procedures. Among those that seem to be given little attention apparently for their low communication harmfulness is the transfer of Polish nasal vowels in contexts where English has a combination of vowel + nasal consonant followed by a fricative. Pronunciation training syllabuses of practical phonetics classes at an academic level and most pronunciation courses or guidebooks are not or are rarely concerned with this phenomenon whereas vowel and diphthong nasalisation accompanied by the deletion of the following nasal consonant remains one of the most persistent errors of Polish learners of English with /ɪ/, /e/, /ə/, /ɑː/, /aʊ/, /ʌ/, /ɒ/ and /əʊ/ most prone to heavy nasalisation and resulting with realisations such as 'consonant' *[kɒ̃sənãt] and 'since' *[sĩs] (Sobkowiak 2004, 204). The fact that the postvocalic nasal consonants pose difficulty for Polish learners of English is based on differences between language systems at a segmental level.

2 Background

The realisation of vowel + nasal consonant sequence in such languages as Polish and English differs considerably in terms of at least one major aspect, namely Polish is traditionally assumed to have distinct nasal vowels in its phoneme inventory that do not depend on the context of the following nasal while in the sound system of English there are no such context-independent nasal vowels. As a result, there are no instances of nasal or nasalised vowels in word-final position. However, the latter are very frequent word-medially preceding nasals, which is a feature common to all human languages, including Polish and English, and resulting from the manner in which they are articulated. The lowering of the velum that has to be involved in the production of a nasal sound and which commences prior to the consonantal closure is a mechanical consequence, not a conscious gesture (as in the case of nasal vowels), of the need to prepare for the articulation of the following nasal sound. This is a type of assimilation that has basic articulatory setting almost identical to the one of nasal vowels. The difference lies in "the duration and extent of the velum lowering" (Sobkowiak 2004, 204). While in the case of nasalised vowels this lowering usually takes place sometime in the middle of the vowel, nasal vowels are characterised by the continuously and consciously lowered velum lasting throughout the whole vowel duration. Additionally, both nasal and nasalised vowels are accompanied by the low nasal timbre which is a consequence of the resonating properties of the nasal cavity (Sobkowiak 2004, 204).

The intricacies of the nasal vowels articulations are often taken for granted by a majority of native speakers of Polish who, equipped with the basic school knowledge of standard Polish, would assume that there are six oral /a/, /ɛ/, /ɔ/, /u/, /i/, /ɨ/[1] and two nasal vowels represented orthographically as 'ę' and 'ą', as in 'węże' [vɛ̃ʒɛ] and 'wąż' [vɔ̃ʃ]. The fact that Polish has distinct nasal vowels that form a separate subgroup in the vowel inventory is a clear but oversimplified statement since there are a lot of complexities surrounding this phenomenon. Nasal vowels are not stable, steady qualities in the way other Polish vowels are. They may differ in terms of strength of nasal resonance, which can be strong or weak and the way it is produced. Nasal resonance can manifest itself as a separate nasal consonant or in some cases it may not be present at all. These complexities are subject to constant scrutiny by the phoneticians and phonologists specialising in Polish linguistics (Dukiewicz 1967; Zagórska-Brooks 1968; Trypućko 1971; Biedrzycki 1978; Wierzchowska 1980; Sawicka 1995—mentioned in Wiśniewski 1998, 77). Their work results with knowledge that is usually absent from school and may seem against lay Polish speakers' intuitions. For instance Wierzchowska (1980) claims that there are not two, but six nasal vowels in Polish:

[ɛ̃] as in 'węże' /vɛ̃ʒɛ/ (snakes)
[ɔ̃] as in 'wąż' /vɔ̃ʃ / (snake)
[ã] as in 'tramwaj' /trãvaj/ (tram)
[ĩ] as in 'instytut' /ĩstɨtut/ (institute)
[ũ] as in 'kunszt' /kũʃt/ (art, craftsmanship)
[ɨ̃] as in 'czynsz' /tʃɨ̃ʃ (rent)

These intuitions may be grounded on the strong attachment to the spelling which does not reflect the realities of [ã], [ĩ], [ũ] and [ɨ̃] in the sense that it does in the case of [ɛ̃] and [ɔ̃]. Interestingly, only these two can be pronounced in a word-final position; others can appear word-medially, which emphasises their dependence on the context and makes their existence less transparent to non-experts.

An important fact that needs to be discussed briefly here is that these vowels are characterised by the asynchronous nasalisation or, in other words, they are of a dual, sometimes diphthongal nature. The first element that they consist of is an oral nasalised segment which does not differ considerably from an equivalent oral vowel, followed by a nasal sound with articulators getting prepared for the next segment. This nasal component may in turn have different realisations either in the form of semivocalic or consonantal nasality depending on the context in which it appears. If a nasal vowel is followed by a plosive or affricate, its second element is a nasal consonant [m], [n̪], [n], [ɲ], [ŋʲ], [ŋ] homorganic to the next sound whose equivalent in spelling appears after letters 'ą' and 'ę':

[1] Throughout this paper IPA transcription is used instead of other systems proposed by the authors quoted here.

- before bilabials:
  ‘tępy’ [tɛ̃mpɨ]
  ‘rąbać’ [rɔ̃mbatɕ]
- before dentals:
  ‘rzędy’ [ʒɛ̃n̪dɨ]
  ‘gorąco’ [gɔrɔ̃n̪tsɔ]
- before alveolars:
  ‘męczyć’ [mɛ̃ntʃɨtɕ]
  ‘rączka’ [rɔ̃ntʃka]
- before prepalatals:
  ‘pędzić’ [pɛ̃ɲdʑitɕ]
  ‘mącić’ [mɔ̃ɲtɕitɕ]
  ‘pięć’ [pʲiɛ̃ɲtɕ]
  ‘piąć’ [pʲiɔ̃ɲtɕ]
- before postpalatals:
  ‘kręgi’ [krɛ̃ŋʲgi]
  ‘łąki’ [wɔ̃ŋʲki]
- before velars:
  ‘męka’ [mɛ̃ŋka]
  ‘urągać’ [urɔ̃ŋgatɕ]

(Ostaszewska, Tambor 2001; 59)

If a nasal vowel is followed by a fricative, its second element is usually a velar semivowel [ɯ̃] produced with a lowered velum: ‘węzeł’ [vɛ̃ɯ̃zew], ‘wąs’ [vɔ̃ɯ̃s].. Before palatal sounds, this element can be pronounced in two ways; it may be either a velar approximant [ɯ̃] whose equivalent traditional symbol in literature on Polish phonetics and phonology is [ɰ̃] or a palatal approximant [j̃]„ both articulated with a nasal resonance:
‘gęś’ [gɛ̃ɯ̃ɕ] or [gɛ̃j̃ɕ]
‘gałąź’ [gawɔ̃ɯ̃ɕ] or [gawɔ̃j̃ɕ]
(Ostaszewska Tambor 2001; 60)

Another peculiarity concerning nasal vowels in Polish is the so-called denasalisation. This phenomenon occurs when instead of a nasal there appears an oral articulation of sounds [ɛ] and [ɔ]. It can take place word-medially before [lʲ] and [w], e.g. ‘wzięli’ [vʑɛlʲi] ‘zaczęła’ [zatʃɛwa], ‘kopnąłem’ [kɔpnɔwɛm] or there may be an optional denasalisation of [ɛ̃] in a word-final position, e.g. ‘drogę’ [drɔgɛ] or [drɔgɛ̃], ‘się’ [ɕɛ] or [ɕɛ̃].

In non-standard pronunciation of Polish, as in many other non-standard language forms, there is a lot of variability. In the case of nasal vowels, the differences in their realisation may depend on the social and geographical conditioning. Within standard Polish itself there exists phonostylistic variation when it comes to foreign borrowings. According to Rubach (1982, 58), it applies to the use of [n] or [ɯ̃] in positions after a vowel and before a fricative:

	GPR[a]	Phonostylistic form
‘instytut’	[ĩns-]	[ĩɯ̃s-] (institute)
‘cenzus’	[tsɛ̃nz-]	[tsɛ̃ɯ̃z-] (qualifications)
‘lansować’	[lãns-]	[lãɯ̃s-](promote)
‘kunszt’	[kũnʃt]	[kũɯ̃ʃt] (art)
‘koncha’	[kɔ̃nxa]	[kɔ̃ɯ̃xa] (conch)

(Rubach 1982, 58)

[a] General Phonetic Representation (GPR)—“a structure reflecting the fact of very slow and monitored speech in opposition to multiple phonostylistic representations, which are the output of phonostylistic rules” (Rubach 1982, 58)

Similar treatment of the issue may be noticed in Wierzchowska (1980, 129) who limits the variation of [ɛ̃n – ɛ̃], [ɔ̃n – ɔ̃] and [ãn – ã] etc. to foreign borrowings only while she accepts the optional realisations of [ɛ̃ – ɛ̃j̃] or [ɔ̃ – ɔ̃j̃] also in the native lexicon of Polish (e.g. 'więź' [vʲɛɕ], 'siąść' [ɕɔ̃ɕtɕ]). Moreover, Wierzchowska states that [ɛ̃], [ɔ̃] and [j̃] are pronounced both in native (e.g. 'kęs' [kɛ̃w̃s], 'mąż' [mɔ̃w̃ʃ], 'koński' [kɔ̃jski]) and borrowed words (e.g. 'sens' [sɛ̃w̃s], 'konsul' [kɔ̃w̃sul], 'w transie' [f trãj̃ɕɛ]) whereas [ã], [ũ], [ĩ], [ỹ] are to be found only in the latter (e.g. 'kamfora' [kãw̃fɔra], 'triumf' [trjũw̃f], 'instytut' [ĩw̃stɨtut] and between words in connected speech ('pan Roman' [pãw̃rɔman).

More recently, Ostaszewska and Tambor (2001, 55) state that the so-called spelling pronunciation of the six sequences of vowels followed by a nasal consonant plus a fricative in borrowings, or Rubach's GPR, is a characteristic of less educated speakers (e.g. [sẽns], [bonãnza])). They emphasise the fact that this way of pronouncing foreign borrowings is incorrect and constitutes an example of hypercorrectness. Yet, they allow certain words to have an optional pronunciation that would be in accordance with their spelling form, taking a prescriptive rather than a descriptive stance, but they do not try to expand on the topic, leaving their readers with mere two examples of such items: [kɔ̃mfort] and [trãmvaj].

Complex articulation of nasal vowels is one of the features that brings about a number of differences between varieties of Polish, both in terms of oral articulation and nasal resonance. In standard Polish their realisation in the form of a vowel + nasal consonant appears always before plosives and affricates while in dialects its distribution may be different. This combination may be found word-medially before fricatives along the eastern border with Belarus (Suwalszczyzna), eastern and north-eastern Mazovian area, as well as parts of southern Silesia, e.g. 'mąż' [mɔnʑ], 'język' [jɨnzɨk] In word-final position the sequence of [ɔ] + nasal consonant is characteristic of the dialects present in the west, centre and south parts of Poland (mainly Great Poland, parts of Mazovian and Silesian areas with neighbouring region of Małopolska), e.g. 'robią' [rɔbʲɔm], 'tą drogą' [tɔm drɔgɔm]. Similarly, it is possible to hear this sequence in the west parts of Małopolska, however with the exception of verbs whose pronunciation may be either [ɔ̃] or denasalised [ɔ] (www.gwarypolskie.uw.edu.pl, see Appendix 1). Interestingly, there is an area in the centre and north of Małopolska where there exists no nasal resonance in any position in a word (Dejna 1981, 1993, Map 41).

3 Research Questions and Design of the Study

The following small-scale study was motivated by the observations of data which was collected for the purpose of a Ph.D. project and constitutes its integral part. Its general aim is to investigate variability in the performance of Polish learners of English in terms of nasalisation. The specific aim is to compare and analyse Polish and Polglish words containing nasal vowels and a combination of vowel + nasal with regard to the general tendencies present in today's Polish language.

3.1 Participants

Fifty 1st year students of English Philology at the University of Łódź took part in the experiment, 42 females and 8 males. They were all recorded in October 2009 at a very beginning of practical phonetics course and prior to any academic pronunciation training. All subjects filled in a questionnaire in which they were asked to state their age, place of birth and place of residence, as well as how long they had been learning English and the frequency and length of any stays in English-speaking countries.

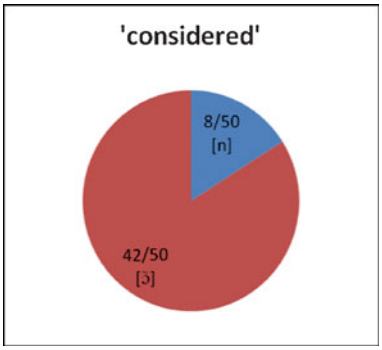
3.2 Data Elicitation

In the course of experiment participants were asked to read two texts, one in English and one in Polish. The former text was a well-known fable *The North Wind and the Sun*, while the latter was a text devised for the purpose of the project and characterised by the equal length, approximate vowel qualities in stressed syllables and overall similar rhythmic structure to *The North Wind and the Sun*. The participants were recorded directly to the PC with the use of MXL Studio 1 USB microphone in not soundproof, but a relatively quiet room.

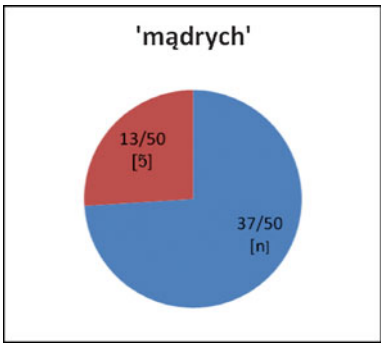
3.3 Measurements

Two words, 'considered' and 'mądrych' were extracted from both English and Polish texts for each subject separately and transferred to Praat. Using manual segmentation techniques the author attempted to detect presence of an alveolar nasal sound /n/ in the spectrographic pictures of both words. It turned out, however, that the instrumental analysis posed problems to the interpretation of acoustic signal in a number of cases as there appeared difficulty in determining distinctive boundaries between a heavily nasalised vowel and a nasal consonant. It then seemed necessary to resort to the ear of the author, not without doubts as well, and eventually both instrumental and auditory analyses were employed in the study.

4 Results and Discussion



Proportions of students pronouncing and not pronouncing a nasal consonant in the word 'considered'



Proportions of students pronouncing and not pronouncing a nasal consonant in the word 'mądrych'

The experiment showed that 42 students (84 %) pronounced the English word 'considered' the Polish way, thus without the closure for the nasal consonant [n] while 37 (74 %)students produced this closure in the Polish word 'mądrych'. One of the assumptions of this study was that if there appears a nasal consonant in 'considered', it should also be present in 'mądrych', since the preferred context for a nasal consonant in Polish is before a stop. It was true for five students. Surprisingly, three students pronounced 'considered' as /kənsɪdɛd/, but 'mądrych' as /mɔ̃drɨx/ (see Appendix 2), which is probably due to spelling pronunciation induced by the type of task that participants performed and/or by the phenomenon of hypercorrection. The difference between the two words proved significant for the tested group.

The above results confirm that transferring a Polish nasal vowel /ɔ̃/ in contexts where English has /ɒn/ or /ən/ followed by a fricative is a frequent phenomenon among Polish learners of English and an element of the Polish accent in English. The results also validate the general tendency of Polish speakers to pronounce a nasal vowel (vowel + nasal approximant) before fricatives and a combination of vowel + nasal consonant before plosives and affricates that seems to be transferred to their target language.

5 Conclusion and Pedagogical Implications

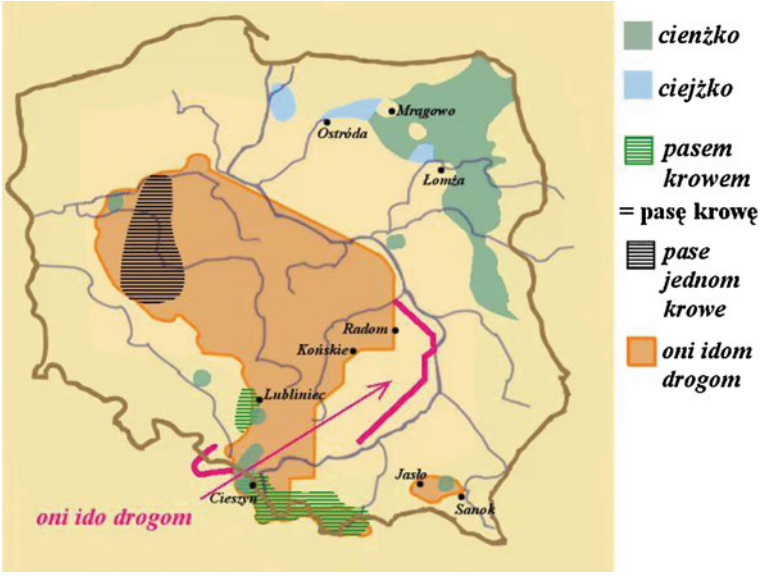
Polish nasal vowels are one of the language features that exhibit variability both within standard Polish, especially as regards foreign borrowings and instances of spelling pronunciation, and between dialects. This variability in realisation of nasal vowels, as the study confirms, exists not only between speakers, but also within an individual. Moreover, the differences present in source and target languages

appear to be an important and intriguing element of the acquisition of the English linguistic system. The issue becomes even more interesting when we realise that the degree of transfer and degree of its persistence may, to some extent, depend on the region one comes from as Polish nasal vowels have a specific regional distribution (Dejna 1993, 195).

A follow-up investigation is required to check whether the reverse choice of lexical items would confirm the tendencies of nasal transfer in English, namely an English word containing /ɒn/ or /ən/ followed by a plosive or fricative consonant and a Polish word with /ɔ̃/ followed by a fricative sound. It should also be investigated whether the students' performance coincides with the places of their origin and/or residence, and whether there exists a correlation between the two. This would require a lot more data and a very controlled way of obtaining it in order to avoid spelling pronunciation.

From the pedagogical point of view, it can be worthwhile to raise awareness among practical phonetics teachers as regards variability in Polish while developing their students' phonological system of English. Bearing in mind some of the tendencies concerning nasal vowels in various contexts and the regional differences between learners' linguistic behaviour, it might prove useful to concentrate on words like 'considered' where a fricative follows rather than treat every vowel + nasal combination as difficult in English.

Appendix 1



Krawczyk-Wieczorek based on Dejna 65, 66, 67, Kucała 15 (www.gwarypolskie.uw.edu.pl).

Appendix 2

The results for 50 participants as for presence (+) or absence (−) of a nasal consonant in words: 'considered' and 'mądrych'.

Speaker	Considered	Mądrych
1	−	+
2	−	+
3	−	+
4	−	+
5	−	+
6	−	+
7	−	−
8	−	+
9	−	+
10	+	−
11	−	−
12	+	+
13	−	+
14	+	+
15	−	+
16	−	+
17	−	−
18	−	+
19	−	+
20	−	+
21	−	+
22	−	+
23	−	+
24	−	+
25	−	−
26	−	+
27	+	+
28	+	+
29	+	−
30	−	+
31	−	+
32	−	+
33	−	−
34	+	+
35	−	+
36	−	+
37	−	+
38	−	−
39	−	+

(continued)

(continued)

Speaker	Considered	Mądrych
40	–	–
41	–	–
42	–	+
43	–	–
44	+	–
45	–	+
46	–	+
47	–	+
48	–	–
49	–	+
50	–	+

References

Based on Karol Dejna 65, 66, 67, Marian Kucała 15. Retrieved 20th November 2010, from <www.gwarypolskie.uw.edu.pl>.

Biedrzycki, Leszek. 1978.*Fonologia angielskich i polskich rezonantów. Porównanie samogłosek oraz spółgłosek*, Warszawa: PWN.

Dejna, Karol. 1993. *Dialekty Polskie*. Wrocław: Ossolineum.

Dejna, Karol. 1981. *Atlas polskich innowacji dialektalnych*. Warszawa—Łódź: Państwowe Wydawnictwo Naukowe.

Dukiewicz, Leokadia . 1967. *Polskie głoski nosowe*. Warszawa: Polska Akademia Nauk.

Krawczyk-Wieczorek, Aleksandra. 2010. *Wymowa rozłożona samogłosek nosowych w gwarach*.

Ostaszewska, Danuta and Tambor, Joanna. 2001. *Fonetyka i fonologia współczesnego języka polskiego*. Warszawa: Wydawnictwo Naukowe PWN.

Rubach, Jerzy. 1982. *Analysis of phonological structures*. Warszawa: PWN.

Sawicka, Irena. 1995. *Gramatyka współczesnego języka polskiego, Fonetyka i fonologia*. Kraków.

Sobkowiak, Włodzimierz. 2004. *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

Trypućko, Józef. 1971. *Polskie samogłoski nosowe w swietle najnowszych badan*. Sprawozdania z Posiedzen PAN, Kraków.

Wiśniewski, Marek. 1998. *Zarys fonetyki i fonologii współczesnego języka polskiego*. Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.

Wierzchowska, Bożena. 1980. *Fonetyka i fonologia języka polskiego*. Wrocław: Ossolineum.

Zagórska-Brooks, Maria. 1968. *Nasal vowels in contemporary standard Polish*. The Hague: Mouton.

[ˈbeɾɚ] or [ˈbetə]? Do Polish Learners of English Accommodate their Pronunciation? A Pilot Study

Magdalena Zając

Abstract One does not have to be a linguist to notice that a person's speech varies with relation to who one is talking to. In the same manner that people often find themselves speaking more slowly and clearly while addressing a foreigner or an elderly person, one will use different language when engaging in casual conversation with friends and when discussing account details with a bank official. To interpret and justify these shifts in communicative behaviour, Communication Accommodation Theory (CAT) may be applied. Although CAT focuses primarily on interactions between native speakers of a given language, there have been several studies which suggest that the theory may also be applied to variation in non-native communicative behaviour (Zuengler 1991). The aim of this pilot study was to investigate whether Polish learners of English accommodate to different English accents. To examine the applicability of CAT in such interactions, a pilot study on 4 first year English philology students was conducted, each of which was interviewed first by an RP and later a Canadian English speaker. The two dependent variables under investigation were tapping and rhoticity. Although the factors which affect the presence or absence of accommodation remain uncertain, the results of this pilot study indicate that some Polish English speakers may accommodate to different accents. It was also observed that tapping appears quite frequently in the speech of the informants selected, which seems to be a result of extensive exposure to American English (tv, cinema, music) rather than an indication of convergence towards any of the interlocutors. The pilot study has also provided some valuable insight into ways of improving the methodology of examining accommodation in native/non-native interactions.

M. Zając (✉)
University of Łódź, Łódź, Poland
e-mail: zajac1234@gmail.com

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_17,

1 Rationale and Aim

Communication Accommodation Theory was developed in the 1970s by Bourhis et al. (1973, in Giles and Ogay 2007) to explain accent and bilingual shifts in interactions, and later evolved into "an interdisciplinary model of relational and identity processes in communicative interaction" (Coupland et al. 1997, 326). The framework is concerned with "predicting and explaining the adjustments individuals make to create, maintain or decrease social distance in an interaction" (ibid.). The two strategies which one can adopt to express attitudes and intentions are termed convergence and divergence. The former consists in making one's speech more similar to the that of the interlocutor so as to communicate social integration or approval, whereas the latter denotes emphasising the speech differences between the interactants so as to manifest social dissociation (ibid.).

Although CAT was originally applied to studying interactions between different linguistic groups or between the native speakers of a given language, there have been several studies which suggest that the model could also be used to examine the communicative behaviour of non-native speakers (Zuengler 1991). For instance, Beebe (1981, in Zuengler 1991) conducted a study on Chinese learners of Thai, who were interviewed in this language by a Chinese and a Thai interlocutor. Beebe (ibid.) examined the participants' production of six Thai vowels and found that five of these vowels were pronounced significantly more Thai-like when the subjects were talking to the Thai interviewer. In other words, the participants seemed to converge towards the Chinese interlocutor by increasing their Chinese accent. Such a phenomenon could be explained by the fact that, according to Beebe and Giles (1984, in Zuengler 1991, 237), "NNSs, because of their nonnativeness, are generally assigned a low-status role when interacting with NSs". Therefore, it could be argued that non-native speakers will converge towards their native interlocutors in an attempt to win their approval.

Since it would appear that CAT can be applied to explain some of the variability in non-native language, it seems worthwhile to investigate whether Polish learners of English accommodate to native interlocutors. Hence, the aim of this pilot study was to examine whether the speakers of Polish will converge their English pronunciation while conversing with representatives of different English accents. To explore this issue, Polish students of English philology were recorded while being interviewed first by a Welsh user of RP, and later by a Canadian English speaker. The subjects were predicted to converge towards the Canadian interlocutor by increasing the amount of tapping in their speech, and to accommodate towards the Welsh interviewer decreasing the amount of rhoticity (as Polish is a rhotic language and non-native speakers may tend to transfer L1 pronunciation patterns into their L2).

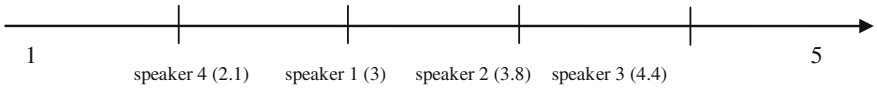


Fig. 1 Level of fluency in English

2 The Study

2.1 Participants

For the purposes of this pilot study, eight conversations with four Polish learners of English were recorded. Each of the informants was recorded twice, first while being interviewed by the Welsh interlocutor, and later while being interviewed by the Canadian interlocutor. The interviewers were two male lecturers from the Institute of English Studies at the University of Lodz. The four interviewees, three females and one male, were first year English philology students from the University of Lodz.

Although the interviewees were first year students, they exhibited different levels of fluency in English. Therefore, 5 experienced teachers of English (Ph.D. students and a Ph.D. from the Institute of English Studies at the University of Lodz) were asked to listen to 2-minute fragments of the recordings and rate fluency in English on a scale from 1 to 5, 1 being very poor fluency and 5 being very good fluency. Fig. 1. shows the participants arranged according to their level of fluency; mean value obtained for each informant is given in brackets. As can be seen, levels of fluency in English are judged to be disparate, speaker 3 being perceived as the most fluent informant and speaker 4 being seen as the least fluent participant.

2.2 Variables

The two native English interlocutors constituted the independent variable in this pilot study. These two speakers were chosen since the accents they use are distinct from each other. We predicted that the participants would use different pronunciation features in order to converge towards the different accents. This way, the degree of accommodation could be measured by investigating whether a participant exhibited more instances of a given pronunciation feature in a conversation with the interlocutor of whose accent this feature is characteristic.

The two dependent variables under investigation were rhoticity and tapping. These two variables were selected partly because the author noticed a lot of variability among Polish learners of English with regard to these features. However, the principal motive for selecting rhoticity and tapping was that they seem to constitute some of the most salient features differentiating the variety of English being learned by the students (which we shall refer to as RP) and a Canadian

English accent. For the purposes of this study, only the speech of the Polish learners was recorded (the interviewers' utterances are audible in the recordings but are not clear enough for analysis). Hence, the RP speaker was assumed to use a non-rhotic accent with little tapping (i.e. using a tap under some circumstances, but only for 't'), while the Canadian interviewer was predicted to exhibit tapping for both 't' and 'd' in a typical post-stress environment and to have a rhotic pronunciation. These assumptions were verified in a subsequent study, in which the two interlocutors talked to ten fourth-year English philology students.

The analysis of the recordings revealed that the Welsh interviewer used a tapped plosive in the place of a/t/in approximately 50 % of the contexts for tapping and was found to use tapping primarily in often-heard words, e.g. in words and phrases such as *but I, that I*, *what is, got a*, *sort of*, *better, little, hated* (which, apart from *hated,* appear in the British National Corpus more than 10 thousand times and as much as 50645 times in the case of *little*), and not in words such as *university, mathematics, dirty, competitive, motivated, frustrating, writing* (which, apart from *university* and *writing,* appear in the BNC less than 3 thousand times). This findings corroborates Shockey (2003) and Wells's (2011) statements that a voiced tap may be heard in Standard Southern British variety, especially in high frequency items such as *but I* or *British*.

In the case of the Canadian speaker, tapping seemed to be much more frequent. Just as the Welsh interlocutor, he was recorded while talking to ten different English philology students during an investigation conducted after this pilot study. Depending on the conversation, the Canadian interlocutor was found to replace/t/ with a tapped plosive either in around 90 % or in 100 % of the contexts for tapping. It was also observed that he used tapping both in content (e.g. *university, tutor, native, meeting, intimidating*) and function words (e.g. *at a, but I, it is, lot of, that a*). Hence, it could be argued that, although present in both interviewer's pronunciation, tapping is a prominent feature only for the Canadian speaker. For this reason, the subjects of this study could be predicted to exhibit more taps when conversing with the Canadian interlocutor.

It is also important to note that, since at the time of the recordings the subjects had just finished their first semester of phonetics classes at the Institute of English Studies, it seems safe to assume that they did not have a very deep knowledge of phonetics. As the majority of English philology students at the University of Lodz are taught to use Received Pronunciation during their phonetics course, and one of the very first issues discussed during such classes is RP's non-rhoticity, the participants are almost certain to have known that a speaker of RP would not pronounce/r/before consonants and pauses. However, whether they would be familiar with the concept of tapping is not as certain, as it is usually discussed when the students gain more experience in phonetics and phonology.

2.3 Procedure

The time of the recorded conversations was around 20 min, the shortest interview lasting for 20 min, and the longest for about 23 min. The topic of the conversations was not specified, since it was assumed that letting the participants talk casually would bear more resemblance to a natural interaction than making them follow a fixed list of questions.

In order to determine whether the subjects converged towards the RP speaker, it was examined whether words containing instances of non-prevocalic/r/were realized as rhotic or non-rhotic. The exact quality of the/r/was not taken into consideration. As for convergence towards the Canadian interlocutor, the subject of the analysis were all the instances of/t/in a post-stress environment preceded by a vowel, liquid or nasal, and followed by a vowel or syllabic/l/. Such a context was chosen since Wells (1986) and Collins and Mees (2008) describe it as one in which tapping might be expected to occur. Also, as in the case of rhoticity, the examination did not concentrate on the exact quality of the tap. If the instance of/t/in the aforementioned environment was voiced, it was counted as a tap. It is alos important to mention that the analysis of the variables was auditory.

3 The Results

3.1 Tapping

For each of the four participants, the proportion of tapping in the conversation with the Canadian interlocutor was contrasted with the proportion of tapping in the conversation with the Welsh interlocutor. The results are provided in Fig. 2. As can be seen, all informants appear to have exhibited some amount of tapping in their pronunciation, regardless of the interviewer. Also, the amount of tapping in the two conversations seems not to vary noticeably for speakers 1, 2 and 4. Nonetheless, a visible difference between the amount of tapping in the two conversations can be observed in the results for speaker 3.

Apart from counting the instances of a given feature, the words in which it occurred were also examined and their type-token ration was calculated (the exact numbers are provided in the appendix to this paper). With regard to speakers 1, 2 and 4, it appears that the words in which they used tapped plosives did not vary much either in relation to the interviewee or the interviewer. The words containing tapped plosives most frequently were *but, that, it* and *lot*. In general, it was observed that the majority of the lexical items in which speakers 1, 2 and 4 exhibited tapping were function words. The frequency of their occurrence was inspected with the aid of the British National Corpus (BNC). It transpired that the greater part of the words in which speakers 1, 2 and 4 used a tapped plosive appear in the BNC more than 20 thousand times. In fact, *but, that* and *lot* appear in this corpus more than 400 thousand

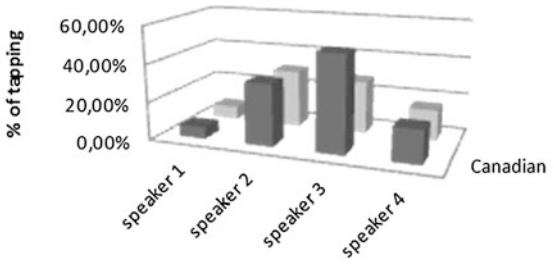


Fig. 2 The total amount of tapping in all contexts for tapping

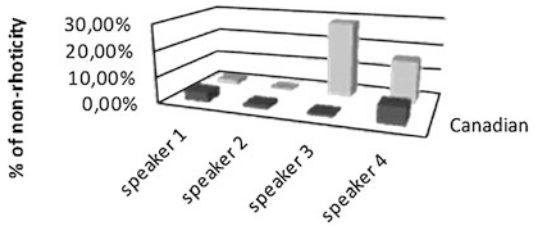
	speaker 1	speaker 2	speaker 3	speaker 4
Canadian	6,10%	32,70%	50,80%	17,50%
Welsh	7,20%	30,60%	28,10%	16,30%

times. There were only three words that seemed to be less frequent, i.e. appear in the BNC less than 20 thousand times. These included *bit*, *account* and *front*. On the whole, it could be argued that the words in which these three participants had tapping were primarily function words which are used frequently in the English language.

As there seemed to be a slight discrepancy between the words in which speaker 3 used tapped plosives in the two conversations, they were calculated separately. The lexical items in which speaker 3 had tapping in the conversation with the Welsh speaker were clearly similar to those in which speakers 1, 2 and 4 exhibited this feature. 5 out of 9 types were function words, and *that* and *but* were among the items which contained tapped plosives most frequently. Also, there was only one word that appears in the BNC less than 20 thousand times, i.e. *hospital*. In other words, as in the case of speakers 1, 2, and 4, speaker 3 seems to have exhibited tapping principally in function words and frequently occurring lexical items during the interview with the Welsh interlocutor. As regards the conversation with the Canadian speaker, *that* was still the most frequent word with tapping. However, only 7 out of 16 types were function words. Moreover, there were as many as 6 words which appear in the BNC less than 20 thousand times (*strategy, automatically, complete, computer, part, thirty*). Hence, it t cannot be said that the words containing tapped plosives in speaker 3's conversation with the Canadian interviewer were primarily those that are used very frequently or that the majority of them were function words.

3.2 Rhoticity

As with tapping, the instances of rhoticity in the two conversations were counted and contrasted for each of the four participants. The results, given in per cent and showing cases where rhoticity did not occur, are shown in Fig. 3. Similarly as in the case of tapping, there is a noticeable difference between the amount of rhoticity in the two conversations in the results for speaker 3. A slighter but visible difference can also be seen in the findings for speaker 4. Nonetheless, the pronunciation of speakers 1 and 2 appears to have been predominantly rhotic during both of the interviews.



	speaker 1	speaker 2	speaker 3	speaker 4
Canadian	3,90%	1,60%	1,00%	6,70%
Welsh	2,10%	0,70%	28,90%	16,20%

Fig. 3 The total amount of non-rhoticity in all appropriate contexts

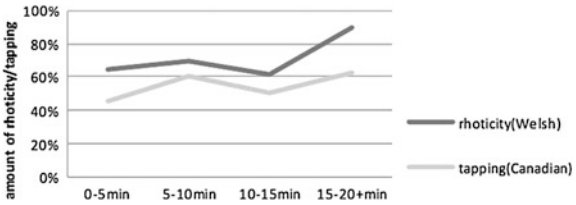


Fig. 4 The amount of rhoticity/tapping over time. Data for speaker 3

3.3 Accommodation Over Time

As speaker 3 appeared to have accommodated towards both of the interlocutors, it seemed worthwhile to investigate whether there was any difference in the amount of accommodation over time. To examine this issue, the two conversations were divided into four five-minute intervals (the last interval began at the fifteenth minute of the recording and might have been slightly longer than five minutes). The amount of rhoticity in the conversation with the Welsh interviewer and the amount of tapping in the interview with the Canadian interlocutor were counted for each of the four intervals and are demonstrated in per cent in Fig. 4. It would appear that both tapping and rhoticity exhibited an upward tendency and their amount increased over time.

4 Conclusions

The study showed that one of the four participants increased the amount of tapping when talking to the Canadian interviewer and decreased the amount on rhoticity when conversing with the RP speaker, while another participant realised a slightly smaller number of words as rhotic when talking to the RP speaker. These findings suggest that some Polish English speakers might adjust their pronunciation towards the accent of their interlocutor.

It was also observed that the only speaker who appeared to converge in both of the conversations was the one considered to be the most fluent in English. At first glance,

this seems to indicate that there might exist a relationship between the degree of accommodation and the level of fluency of a non-native speaker Nevertheless, speaker 2, who was perceived to be almost as fluent as speaker 3, seems not to have accommodated, neither to the Welsh interlocutor nor the Canadian one, whereas speaker 4, who was considered to be the least fluent participant in this pilot study, decreased the amount of rhoticity in the interview with the RP speaker. With regard to speaker 2, his lack of accommodation could arguably be ascribed to the fact that, according to Giles and Ogay (2007), men may be less prone to accommodate than women, and speaker 2 is the only male participant in this study. No explanations have been found as to why Speaker 4 was seen to be accommodating even though her fluency in English was considered to be relatively poor. Hence, whether there might exist a relationship between the degree of accommodation and the level of fluency of a non-native speaker remains uncertain.

Also, the fact that the amount of tapping in speaker 3's interview with the Canadian interlocutor and the amount of rhoticity in the conversation with the Welsh interviewer grew over time may indicate that the latter feature was adjusted consciously, while the former was not. It seems as if at the beginning of the first interview, speaker 3 realised that the interlocutor was using RP and attempted to converge towards him, but as time passed she grew more relaxed and began to forget to pronounce certain words without rhoticity. Since speaker 3 showed a greater amount of tapping towards the end of the conversation with the Canadian interlocutor, it might signify she was not aware that it is one of the features of a Canadian accent and began to use it subconsciously as she became accustomed to this interviewers pronunciation. This might be possible especially since the participants could be assumed to be better acquainted with the concept of rhoticity than with the concept of tapping (see Sect. 2.2). On the whole, this observation suggests that accommodation in the case of non-native speakers may operate both at a conscious and subconscious level.

It was also observed that all the participants exhibited some amount of tapping irrespective of their interviewer. This, together with the fact that the majority of the instances of this feature occurred in function and frequently used words, might indicate that the use of tapping by Polish learners of English is a result of extensive exposure to American English, especially since the preponderance of contemporary movies, TV series and songs are performed in American accents of English, in which tapping appears to be more prominent than in British English (Shockey 2003).

Moreover, it was observed that speaker 3 used tapped plosives in content words more frequently during the conversation with the Canadian interlocutor than she did with the Welsh interlocutor. This finding seems to provide additional support to the claim that speaker 3 was accommodating towards the Canadian native.

One must bear in mind that all the aforementioned conclusions are very tentative and preliminary since the number of participants in this study was very limited and a significantly larger group of subjects is necessary to obtain conclusive results. Nonetheless, this pilot study suggests that accommodation is possible in the case of non-native speakers, and that although it appears to be a highly complex phenomenon, it seems to be worth studying further. Apart from that, this pilot study provided some valuable insight into the ways of improving the methodology of such a research

project. A follow-up study should be conducted on a larger group of participants, which ought to include an even amount of representatives of the two genders. The topic of the conversations could be fixed, so that the informants would use relatively similar words in both conversations. Hence, a semi-structured interview might be better than letting the participants have a completely casual conversation. For instance, the Map Task could be used. Recording the subjects before they talk to the native speakers could also be useful. Such recordings could serve as a point of reference when examining the informants potential convergence or divergence. Also, the native interviewers need to be recorded together with the interviewees so that one may ascertain whether the features one is investigating are present or absent in the natives' pronunciation and whether the native speaker used a given feature in a particular conversation. Another important change in the methodology should be to include spectrographic analysis in the examination of the recordings so as to gain a more objective assessment of the results. Finally, it would surely prove interesting to examine features other than rhoticity and tapping.

Appendix

Table A.1 The words containing tapped plosives for speakers 1, 2 and 4

Type	Token
But	18
That	8
It	7
Lot	6
Get	3
Not	2
What	2
About	2
Out	2
Bit	2
At	1
Better	1
Yet	1
Cut	1
Cannot	1
Account	1
Front	1
Little	1
18	60
TTR	30

Table A.2 The words containing tapped plosives for speaker 3, conversation with the Canadian interlocutor

Type	Token
That	4
Start	2
But	2
What	2
Hospital	2
Not	1
Little	1
Better	1
About	1
9	16
TTR	56,3

Table A.3 The words containing tapped plosives for speaker 3, conversation with the Welsh interlocutor

Type	Token
That	6
Better	5
But	3
Water	3
City	2
Strategy	2
What	2
About	1
At	1
Automatically	1
Complete	1
Computer	1
Lot	1
Part	1
Sort	1
Thirty	1
16	32
TTR	50

References

Beebe, Leslie. 1981. Social and situational factors affecting the strategy of dialect code-switching. *International Journal of the Sociology of Language* 32:139–149

Beebe, Leslie and Giles, Howard. 1984. Speech accommodation theories: a discussion in terms of second-language acquisition. *International Journal of the Sociology of Language* 46:5–32

Bourhis, Richard, Giles, Howard and Taylor, Donald. 1973. Towards a theory of interpersonal accommodation through language: some Canadian data. *Language in Society* 2:177–192

Collins, Beverly and Mees, Inger. 2008. *Practical phonetics and phonology. A resource book for students*. New York: Routledge

Coupland, Nikolas and Jaworski, Adam. 1997. Relevance, accommodation and conversation: Modeling the social dimension of communication. *Mutlilingua* 16:235–258

Giles, Howard and Ogay, Tania. 2007. Communication Accommodation Theory. In Wendy Samter and Bryan Whaley, (eds) *Explaining communication. Contemporary theories and exemplars.* New Jersey: Lawrence Erlbaum Associates, Inc.: 325–244

Skockey, Linda. 2003. *Sound patterns of spoken English*. Oxford: Blackwell Publishing Ltd

Wells, John Christopher. 1986. *Accents of English.* Cambridge: Cambridge University Press

Wells, John Christopher. 2011. details. John Wells's phonetic blog. (http://phonetic-blog.blogspot.com/2011/09/details.html). [Last accessed 22nd February 2012]

Zuengler, Jane. 1991. Accommodation in native-nonnative interactions: Going beyond the "what" to the "why" in second language research. In Justine Coupland, Nikolas Coupland and Howard Giles (eds) *Contexts of accommodation: developments in applied sociolinguistics*. Cambridge: Cambridge University Press: 223–245

Glottal Stops in General American (Intervocalic Environments)

Przemysław Ostalski

Abstract This paper aims at investigating the occurrence of glottal stops in General American accents in intervocalic phonetic environments. So far this particular aspect of American pronunciation has only received occasional mention in the literature, especially when compared to various descriptions and analyses of the use of glottal stops in Britain. The present study focuses primarily on the age factor and whether or not it has any significant bearing on the frequency of glottal stops. In this comparative pilot study, three speakers were investigated using selected samples: Andy Rooney (age 91), Keith Malley (age 36) and Jesse Thorn (age 29). The speakers use moderately standard form of General American and what may be classified as a semi-casual style, being somewhat monitored and thus falling in between interview style (IS) and casual style (CS) in Labovian terms. The study reveals interesting phonetic conditioning for glottal stops especially in the intervocalic position. Certain vowels preceding or following the consonant /t/ seem to trigger glottalling more readily than others.

1 Introduction

Glottal stop is a common allophone of /t/ in various phonetic environments in most varieties of English. However, their frequency appears to be higher in e.g. many English English accents, Scottish English especially as compared to their frequency in General American. Thus, glottalling on the whole is not considered to be a prominent feature of American accents, this is especially true of intervocalic

P. Ostalski (✉)
University of Łódź, Łódź, Poland
e-mail: ostalski@uni.lodz.pl

E. Waniek-Klimczak and L. R. Shockey (eds.), *Teaching and Researching English Accents in Native and Non-native Speakers*, Second Language Learning and Teaching,
DOI: 10.1007/978-3-642-24019-5_18,

environments. Pronunciations such as *a lo*[ʔ] *of*, *be*[ʔ]*er* or *ci*[ʔ]*y* are, therefore, usually viewed as characteristic of e.g. Cockney but not of General American, where voiced taps are common realizations.

The literature describing glottalling of /t/ (and also remaining voiceless plosives) in various phonetic environments in various English English varieties is quite abundant. Descriptive perspectives may range from purely phonetic, through phonostylistic to sociolinguistic (see e.g. Coggle 1993; Cruttenden 1994; Fabricius 2000; Milroy et al. 1994; Wells 1982, 1997).

Our investigation of t-glottalling in General American was prompted by the statement by that glottal stops do appear in intervocalic environments in General American, there being a clear preference for them in younger speakers. The essential goal of this study is to explore the way age factors and phonetic environments influence t-glottalling in General American. We hope that our analysis will contribute in a significant, if small way to forming a clearer picture of glottalling phenomena in American accents.

2 Previous Studies of t-Glottalling in American Accents

As was stated above, there are only few studies describing glottalling in American accents. Before we describe our analysis, however, a brief survey of glottalling studies of American accents that do exist will be presented in this section so that our results might later be cross compared.

2.1 J. C. Wells (1982, 2008, http://www.phonetic-blog.blogspot.com/)

Wells only mentions in passing possible phonetic environments where glottal stops are frequent realizations of /t/ in General American (he also describes extensive glottalling in Appalachian accents, these, however, clearly fall outside General American):

1. _# obstruent (*quite good, it can, that pan*)
2. _# liquid or semivowel (*it was, quite likely*)

The first two environments contain alveolar gesture accompanying the glottal closure—preglottalized coarticulation [ʔ͡t], rather than just a glottal stop.

3. _ syllabic nasal (*button, sentence, Clinton*)

Glottalling (glottal replacement) seems to be obligatory in most kinds of General American and as pointed out by Wells only accents from Boston area use flapping in this environment (e.g. [ˈbʌt̬n̩]). Additionally the cluster /Vntn̩/ is

typically pronounced as nasalization + [ʔn̩], thus *sentence* [ˈsɛ̃ʔn̩s], *Bill Clinton* [bɪɬ ˈkl̃ĩʔn̩]. Noteworthy is the fact that there is no mention of possible glottalling in intervocalic environments.

2.2 J. Roberts (2006) As Old Becomes New: Glottalization in Vermont

In this study 47 Vermonters from the northwestern rural region, aged 3–80 were interviewed. Only the age and sex dimensions were investigated. The study found that younger speakers glottalized more, men favored glottal stops more than women and finally that /Vt#V/ was the only phonetic environment in which glottal stops and flaps alternated.

2.3 D. Eddington and Taylor (2009) T-Glottalization in American English

This particular study used, among other techniques, shadowing technique, where participants hear an utterance and immediately repeat it several times. However, the crucial part of the utterance they hear has been masked with noise so that the participants are not merely mimicking what they hear, but producing utterances uninfluenced by the pronunciation in the presented speech. One might point out a potential drawback of this particular technique of elicitation, namely that it draws attention to the environment in question (natural self-monitoring).

The study involved 58 participants (27 women, 31 men), with college education, aged 19–49, from various states (they were grouped as only Western vs. non-Western) and 20 collocations with the environment of /Vt#V/ were used.

The results showed that only 10 % of the tokens were glottal stops, also that front vowels following /t/ favored glottalling and finally that glottalling was also favored by women, younger speakers and Westerners (California).

Furthermore, in the shadowing experiment, the 20-year-old females produced glottal stops in 20.4 % of the cases. In the reading task this increased to a staggering 55.8 %. This high degree of glottalling in a task that tends to elicit very formal language clearly demonstrated that the use of the glottal stop was not (socially) stigmatized, which fact is in stark contrast to the sociolinguistic situation concerning glottalling obtaining in e.g. England, where the intervocalic environment seems to be particularly stigmatized (Wells 1982).

2.4 D. Eddington and Channer (2010) American English Has Go? A Lo? of Glottal Stops: Social Diffusion and Linguistic Motivation

This study analyzed 1,101 instances of word-final, prevocalic /t/ from the Santa Barbara Corpus. The glottal stop occurred in 24 % of the cases. Factors that favored glottaling of /t/ were: age and region (young speakers and Westerners) and a following stressed syllable. However, gender and following vowel quality were not influencing variables.

The authors also pose an interesting question, specifically: if prevocalic glottaling is uncommon word-internally in American English, why is it apparently spreading word-finally? Their answer is that word-final /t/'s are more often followed by word-initial consonants than vowels, which places them in a glottalizing context. Instances with a glottal realization are therefore stored in the mental lexicon and are available as possible pronunciation choices even in prevocalic position.

It is clear then that glottal stops appear in intervocalic positions in American accents indeed. They are seemingly quite frequent word-finally (*got it*), not so much word-internally (*city*). There are also two categories of factors that typically influence the frequency of glottalling: extralinguistic (age, gender, region, style) and linguistic (preceding and/or following vowel quality, presence or absence of stress in the preceding or following syllable, word/collocation frequency). This last purely phonetic category seems to be uniquely American in the sense that in all the remaining glottalizing regions of the English speaking world only extralinguistic factors play a significant role in the frequency of glottalling (Wells 1982).

Our analysis is meant to investigate this possible interaction of both extralinguistic and purely linguistic factors, and the frequency of glottalling in General American. The analysis concentrates only on intervocalic environments, as they are traditionally considered to constitute a genuine English-American difference (glottal or alveolar stops vs. voiced alveolar taps) and the occurrence of glottal stops in these particular environments in American accents is certainly unexpected. Since our analysis is but a humble contribution to an ongoing investigation of phonetic and phonological variability found in General American, only selected extralinguistic variables were chosen and others were held stable.

3 The Analysis of Glottalling in Three General American Speakers

This section describes the analysis concerning glottalling in three speakers of General American. First we will briefly describe the procedure, then the subjects, variables, general and particular results will be described.

3.1 The Procedure

This study relied primarily on impressionistic analyses, although we were prepared to use any acoustic software (e.g. Praat) for unclear cases. There are a number of reasons behind this particular choice of procedural steps. First, as only intervocalic environments were investigated (VtV), the difference between a glottal stop and a tap was deemed to be highly detectable auditorily. Then, the analysis was purely binary, that is, it only focused on the choice: a glottal stop vs. some other sound, and possible realization that were not actually glottal stops (tap voiced or voiceless, plosive) were disregarded. It has to be admitted though, that in most cases it was in fact a voiced tap that was used in environments in question. Finally both the high quality of recorded materials and partly the chosen style allowed for a sufficient auditory analysis.

3.2 The Subjects and Materials

As was mentioned above, three speakers of General American with a similar background and profession were chosen for the purposes of the analysis. Below we present a short description of the subjects:

1. Andrew Aitken "Andy" Rooney (born January 14, 1919) aged 91, born in Albany, New York, most notable for his weekly broadcast "A Few Minutes with Andy Rooney", a part of the CBS News program 60 Minutes since 1978. college education
   3 h of his speech were analyzed (90 samples from 2009-07 till 2010-09) downloaded from http://www.cbsnews.com
2. Keith Malley (born April 15, 1974) aged 36, born in Somerset, Pennsylvania; has lived for 20 years in and around New York City and is a stand-up comedian and podcaster who founded the Keith and The Girl podcast.
   college education
   7 h of his speech were analyzed (4 podcasts, 2010-08-10 through 08-15) downloaded from http://www.keithandthegirl.com/
   The difference in the number of hours analyzed in the case of Rooney and Malley stems from the fact that Rooney's broadcast is mainly a monologue, whereas Malley is frequently involved in conversations with other people.
3. 3. Jesse Thorn (born April 24, 1982), aged 28, born in Bridgeport, Connecticut; has lived in New York, San Francisco and Los Angeles and is the host and producer of the radio show and podcast The Sound of Young America
   college education
   5 h of his speech were analyzed (6 shows, 2010-09) downloaded from http://www.maximumfun.org/shows/

3.3 The Summary of Variables

Due to the fact that this particular analysis was meant as only preliminary and was carried out more to confirm the results of previous studies than to discover new conditioning of glottalling in General American, only one essential extralinguistic variable was chosen with the remaining variables being stable.

Controlled (stable) variables applicable to the analysis were: gender (male), variety (non-western GenAm—Labov et al. 2005), style (in between interview style (IS) and casual style (CS) in Labovian terms, perhaps best described as pseudo-formal), rate of delivery (mean 170.9 words per minute; Rooney's 170.1 wpm, Malley's 166.9 wpm, Thorn's 175.6 wpm; cf. 150 wpm moderate recommended for audiobooks, thus the tempo used by the three speakers was not allegro; moderato perhaps), education (college), profession (show-hosts) and social class (middle class).

There was only one true independent variable (the other being potentially the quality of preceding and following vowels). The variable in question was age of the speakers. The speakers were aged 91, 36 and 28 years old. A huge gap between 91 and 36 was included on purpose in the analysis, as the previous studies showed a clear preference for glottalling in intervocalic environments in younger speakers, this particular age-gap was expected to be reflected as a sudden, marked increase in the frequency of glottal stops in this specific phonetic environment.

Finally, as mentioned earlier the dependent variable in the analysis was glottalling of /t/ in the intervocalic environment in General American. As the results of the previous studies showed that glottalling was frequent word-finally and infrequent or virtually absent word-medially, it was decided to include only the former (/Vt#V/) in the analysis.

3.4 General Results

There were 446 tokens of /Vt#V/ in the material analyzed, of which only 9 % contained a glottal stop. This result is quite surprising, considering a substantial frequency of glottalling in General American as revealed by the previous studies. A summary of general results is presented in Table 1.

An even more interesting fact is revealed when we look at the correlation of age and the frequency of glottalling. Recall, that on the basis of the previous studies a negative correlation (an increase in age matches a decrease in glottalling) is to be expected. When Pearson correlation coefficient is computed, however, we obtain $r = 0.42988$. Thus, first of all, the correlation is very weak, we could even go as far as saying that it is nonexistent. Second, the correlation is positive which is quite unexpected. We can therefore conclude that there seems to be no important correlation of age and glottalling (in these males) and age is not a significant factor influencing the frequency of glottalling in our sample. Obviously, this tentative

Table 1 Glottalling per speaker, number of tokens

Subject	Age	Glott.%	Glott. N	N
Rooney	91	9,220	13	141
Malley	36	9,494	15	158
Thorn	28	8,163	12	147
Mean	51.667	8.969	13	148.667
Total			40	446

conclusion could be challenged on the basis of the fact that the sample analyzed is quite small. Were the correlation to be in actual fact significant in General American on the whole, however it should be expected to be found in any randomly drawn sample (as ours is) regardless of its size.

Additionally (single factor) ANOVA was applied to the data to confirm this apparent lack of correlation of age and glottalling. The reasoning was as follows: given, for the sake of argument, the null hypothesis (H_0)—that glottalling in General American is not dependent upon the age of speakers, the p-value of observations (glottalling per speakers and per phonetic environment—see Appendix I, it should be noted that vowel transcriptions are phonemic not phonetic) should exceed 0.05 level of significance. The calculated p-value for the sample was in fact 0.9, so the observation is fairly consistent with the null hypothesis, which cannot be rejected. Thus, the difference in the frequency of glottalling among the three speakers is not statistically significant, perhaps reflecting random variation. If age should play any role in glottalling, then the difference ought to be statistically significant. We have obtained thus yet another confirmation that age seems not to influence glottalling and as explained earlier any argument concerning the size of the sample cannot in actuality be made here.

3.5 Specific Results

Having investigated the influence of age upon glottalling, or in fact the lack thereof, we go on to the discussion of the influence of preceding and following vowels on the frequency of glottalling. As in the previous case ANOVA (two way with replication) was applied to the data (glottalling per preceding and following vocalic environment—see Appendix II) to see whether preceding or following vowels influenced in a statistically significant way the frequency of glottalling. The p-value for the rows (following vowels) was 0.00002587 and the one for columns (preceding vowels) was 0.0034635. With the significance level of 0.05, both environments interact with glottalling in a statistically significant way. With both p-values being seriously lower than the significance level, the null hypotheses (1. the quality of a preceding vowel does not influence the frequency of glottalling; 2. the quality of a following vowel does not influence the frequency of glottalling)

Table 2 Correlation of preceding vowel advancement/height and glottalling

	Vowel advancement			Vowel height	
	Degree of a.	Glott. %		Degree of h.	Glott. %
Front	3	85	High	3	92.5
Central	2	11	Mid	2	7.5
Back	1	4	Low	1	0
Degree of a.	1		Degree of h.	1	
Glott. %	0.90160396	1	Glott. %	0.90021045	1

must be rejected. Thus we can safely conclude that the exact frequency of glottalling is dependent upon the quality of surrounding vowels.

Now, if we group the vowels according to their height and advancement and compute the *r* coefficient, some remarkable tendencies emerge.

It is clear from Tables 2 and 3 that both preceding front and high vowels favor glottalling (high positive correlation). As for the following vowel, the situation does not seem to be clear-cut at first glance, however if the oppositions mid-non-mid and central-non-central are introduced instead of three-way distinctions, a clearer picture emerges with central and mid vowels favoring glottalling.

Although there does appear to exist a set of allophonic rules predicting glottalling in General American, it is extremely difficult to find a plausible phonetic motivation behind them. A prototypical phonetic environment for word-final intervocalic glottalling in our data seems to be /iːt#ə/ or /ɪt#ə/ and a tricky question is what actually causes so frequent a glottalling rate phonetically speaking.

A surprising answer may be found, however, when instead of focusing on phonetic motivation only, we shift our attention in the direction of sound frequency itself.

Table 4 displays the frequency of glottalling before and after certain vowels in our study together with sound frequency taken from Knowles (and reprinted in Cruttenden). Supposing now, that glottalling is in fact allophonically tied with particular vowels (high, front, mid etc.), then we would clearly expect a significant statistical difference among the three columns. Conversely, if we find that there is no statistically significant difference among the three columns, we can safely conclude that what at first glance appears to be a purely allophonic rule is in fact only governed by sound frequency. As it turns out, (single factor) ANOVA yields the p-value 0.998 (at the significance level 0.05), which entails no statistically significant difference. Simply put then, glottalling is frequent after or before certain vocalic qualities not because there is something intrinsically extraordinary about the vowels themselves but because the vowels in question are more frequent. A real life analogy somewhat jocular for this would be the difference in obesity between adolescents and grown-ups (after Gordon-Larsen P., L. S. Adair, M. C. Nelson and B. M. Popkin). To say that there are more obese grown-ups than adolescents is stating the obvious. To attribute this difference to the discrepancy in the amount of food consumed, however, would be overly simplistic, as the fact

Table 3 Correlation of following vowel advancement/height and glottalling

	Vowel advancement			Vowel height	
	Degree of a.	Glott. %		Degree of h.	Glott. %
Front	3	11	High	3	15
Central	2	75	Mid	2	80
Back	1	14	Low	1	5
Degree of a.	1		Degree of h.	1	
Glott. %	0.40160396	1	Glott. %	0.29002104	1

Table 4 Glottalling frequency per preceding and following environment and sound frequency

	Vt	tV	*f* (Knowles 1987)
i	30	5	4.20811
ɪ	27.5	5	21.24458
ɛ	7.5	0	7.574598
æ	0	0	3.698036
ʌ/ə	0	55	31.85412
ɑ	0	0	5.508799
ɔ	0	0	3.162459
ʊ	7.5	0	2.193318
u	7.5	0	2.881918
aɪ	0	17.5	4.667177
ɔɪ	0	0	0.357052
eɪ	10	15	4.361132
ʊ	7.5	0	1.555726
oʊ	2.5	2.5	3.851058

remains that grown-ups tend to be obese more frequently because they tend to move a lot less frequently than adolescents, not because they eat a whole lot more.

4 Conclusions

The analysis, although only preliminary showed no marked connection between age in glottalling in General American. It also showed a possible interaction of sound frequency and the frequency of glottalling. Clearly, glottallization should be researched in greater detail, we can offer, however, certain reasonable conclusions. Glottalling in General American appears to be a low-level allophonic rule affecting /t/ in intervocalic environments occasionally and somewhat inconsistently in a very small sample from a specific location. Additionally it is intimately tied with sound frequency. A lot more remains to be discovered about the issue at hand undoubtedly. A working hypothesis is that glottalling in intervocalic environments in General American is as frequent (or infrequent indeed) as is tapping in exactly the same environment in, say, RP.

Appendix I

Environ.	Rooney	Malley	Thorn
i-	3	6	3
ɪ-	4	4	3
ɛ-	1	1	1
æ-	0	0	0
ʌ/ə-	0	0	0
ɑ-	0	0	0
ɔ-	0	0	0
ʊ-	1	1	1
u-	1	1	1
ɚ-	0	0	0
ɝ-	0	0	0
aɪ-	0	0	0
ɔɪ-	0	0	0
eɪ-	2	1	1
aʊ-	1	1	1
oʊ-	0	0	1
-i	1	0	1
-ɪ	0	1	1
-ɛ	0	0	0
-æ	0	0	0
-ʌ/ə	7	9	6
-ɑ	0	0	0
-ɔ	0	0	0
-ʊ	0	0	0
-u	0	0	0
-ɚ	0	0	0
-ɝ	0	0	0
-aɪ	3	2	2
-ɔɪ	0	0	0
-eɪ	1	3	2
-aʊ	0	0	0
-oʊ	1	0	0

Appendix II

	i-	ɪ-	ɛ-	æ-	ʌ/ə-	ɑ-	ɔ-	ʊ-	u-	ɚ-	ɝ-	aɪ-	ɔɪ-	eɪ-	aʊ-	oʊ-
-i	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-ɪ	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-ɛ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-æ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-ʌ/ə	5	6	2	0	0	0	0	2	2	0	0	0	0	2	2	1
-ɑ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-ɔ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
(-ʊ)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-u	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-ɚ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-ɝ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-aɪ	2	3	0	0	0	0	0	0	1	0	0	0	0	1	0	0
(-ɔɪ)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-eɪ	1	1	1	0	0	0	0	1	0	0	0	0	0	1	1	0
-aʊ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
-oʊ	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

References

Coggle Peter. 1993. Do you speak Estuary? London: Bloomsbury.

Cruttenden Allen. 1994. Gimson's pronunciation of English. London: Edward Arnold.

Eddington Daniel and Cindy Channer. 2010. *American English has go? a lo? of glottal stops: Social diffusion and linguistic motivation*. American Speech 2010 85(3):338–351.

Eddington Daniel and Martin Taylor. 2009 T-glottalization in American English. American Speech 2009 84(3):298–314.

Fabricius Andrus 2000. T-glottaling between stigma and prestige: A sociolinguistic study of Modern RP. Unpublished Dissertation: Copenhagen Business School.

Labov William Ash Samuel and Boberg Carl 2005. The atlas of North American English. Berlin: Mouton de Gruyter.

Milroy John Milroy. Linsey Hartley Sam and Walshaw Donald 1994. Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. Language Variation and Change. 6. 327–357.

Roberts Julie. 2006. As old becomes new: Glottalization in Vermont. American Speech . 81. 227–249.

Wells John C. 1982. Accents of English. vol. 3. Cambridge: Cambridge University Press.

Wells John C. 1997. Whatever happened to received pronunciation? In C. Medina Casado and C. Soto Palomo eds. II Jornadas de Estudios Ingleses . pp. 19–28. Jaen: Universidad de Jaen.

Wells John C., 2008. Longman Pronunciation Dictionary. Third edition with CD. Harlow: Pearson Education Limited.