

Educational Linguistics

Dongbo Zhang
Ryan T. Miller *Editors*

Crossing Boundaries in Researching, Understanding, and Improving Language Education

Essays in Honor of G. Richard Tucker



Springer

Educational Linguistics

Volume 58

Series Editor

Francis M. Hult, Dept. of Education, Sherman Hall A Wing,
University of Maryland, Baltimore, MD, USA

Editorial Board Members

Marilda C. Cavalcanti, Universidade Estadual de Campinas, Campinas, Brazil

Jasone Cenoz, University of the Basque Country, Leioa, Spain

Angela Creese, University of Stirling, Stirling, UK

Ingrid Gogolin, University of Hamburg, Hamburg, Germany

Christine Hélot, Université de Strasbourg, Strasbourg, France

Hilary Janks, University of the Witwatersrand, Johannesburg, South Africa

Claire Kramsch, University of California, Berkeley, USA

Constant Leung, King's College London, London, UK

Angel Lin, Simon Fraser University, Burnaby, Canada

Alastair Pennycook, University of Technology, Sydney, Australia

Educational Linguistics is dedicated to innovative studies of language use and language learning. The series is based on the idea that there is a need for studies that break barriers. Accordingly, it provides a space for research that crosses traditional disciplinary, theoretical, and/or methodological boundaries in ways that advance knowledge about language (in) education. The series focuses on critical and contextualized work that offers alternatives to current approaches as well as practical, substantive ways forward. Contributions explore the dynamic and multi-layered nature of theory-practice relationships, creative applications of linguistic and symbolic resources, individual and societal considerations, and diverse social spaces related to language learning.

The series publishes in-depth studies of educational innovation in contexts throughout the world: issues of linguistic equity and diversity; educational language policy; revalorization of indigenous languages; socially responsible (additional) language teaching; language assessment; first- and additional language literacy; language teacher education; language development and socialization in non-traditional settings; the integration of language across academic subjects; language and technology; and other relevant topics.

The *Educational Linguistics* series invites authors to contact the general editor with suggestions and/or proposals for new monographs or edited volumes. For more information, please contact the Editor: Marianna Georgouli, Van Godewijckstraat 30, 3300 AA Dordrecht, The Netherlands.

All proposals and manuscripts submitted to the Series will undergo at least two rounds of external peer review.

This series is indexed in Scopus and the Norwegian Register for Scientific Journals, Series and Publishers (NSD).

Dongbo Zhang • Ryan T. Miller
Editors

Crossing Boundaries in Researching, Understanding, and Improving Language Education

Essays in Honor of G. Richard Tucker

 Springer

Editors

Dongbo Zhang 
School of Education
University of Exeter
Exeter, UK

Ryan T. Miller 
Department of English
Kent State University
Kent, OH, USA

ISSN 1572-0292

ISSN 2215-1656 (electronic)

Educational Linguistics

ISBN 978-3-031-24077-5

ISBN 978-3-031-24078-2 (eBook)

<https://doi.org/10.1007/978-3-031-24078-2>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

Foreword

It is my honor to have been asked to provide a foreword to this volume, a tribute to the distinguished career of G. Richard (Dick) Tucker on the occasion of his 80th birthday. For over 50 years, Dick has inspired the profession with his leadership and keen sense of what it means to realize innovations in the field and dispel artificial borders that impeded the cultivation of new knowledge. Dick's influential research is far-reaching and ranges from early psycholinguistic studies on various aspects of bilingualism to later work in sociolinguistics, heritage language learners, and implementation and assessment of innovative foreign language programs in the United States.

The title of this volume is apt. By crossing boundaries, Dick has significantly contributed to research in diverse domains within the field of applied linguistics. He has forged a pathway for professional practice by revealing to the applied linguistics community how to link what seems impossible, disparate, and incompatible. All the studies in this volume are inspired and shaped by Dick's extensive contributions to the field of applied linguistics. They address the realities of becoming bilingual, the critically important role of foreign language teacher education, the process of innovative language program design and assessment, and policy and planning for literacy development among minority and ethnic populations around the world.

My association with Dick began long before I started regularly working with him in the early 1990s on foreign language in elementary schools and on mentoring doctoral students in second language acquisition and foreign language education programs. My first encounter with Dick occurred in 1986, when I was a graduate student at the University of Delaware serving as a teaching assistant in the Department of Modern Languages. At that time, the department was undergoing an external evaluation. I was told by the department chair that my intermediate French class, taught using DiPietro's Strategic Interaction approach, would be observed by the evaluation team. Little did I know that the team would consist of three notables in the field: Adam Makai, Earl Stevick, and *Dick Tucker*.

I vividly remember this observation by the three-member evaluation team. They sat on the right side of the classroom while students participated in a lesson on rules of the road in French. Little did I know I would cross paths with Dick again in 1992

when he would become Professor of Applied Linguistics in the Department of Modern Languages at Carnegie Mellon University (CMU) in Pittsburgh, PA. In 1995, Dick became department head, a position he held until 2007. In 2008-2009, he was named interim Dean of Student Affairs and interim Dean of CMU Qatar in 2010–2011. After his year in Qatar, he returned to the Modern Languages faculty until his retirement in June 2015. Dick's upward trajectory of leadership roles at CMU attests to his efficiency and organizational skills as an administrator and academic leader.

While responding to the demands of university administration, Dick remained a prolific scholar and researcher. His research reflects two areas for which he is well known as an international authority – namely, language policy and planning and innovative second and foreign-language program development and evaluation. While serving as department head at Carnegie Mellon University, Dick continued to write extensively on language policy and practice, analyzing issues such as the English-only movement, the education of linguistically and culturally diverse students, and the need for a language competent citizenry.

During his tenure at CMU, Dick and I established a productive professional relationship while working on implementing, monitoring, and evaluating innovative Japanese and Spanish elementary and middle school foreign language programs (see for example, Tucker & Donato, 2003). During these years, Dick was masterful at building cooperative relationships with a dedicated group of doctoral students at CMU and the University of Pittsburgh who, across several generations, served as active members of our research team, participants in the elementary school language classrooms that we investigated, and co-authors of many publications about our findings (see, for example, Chinen et al., 2003; Dominguez et al., 2005; Donato et al., 1996, 2015; Tucker et al., 1996, 2001).

To facilitate our work. Dick forged relationships with school district administrators and teachers who gave us access to real classrooms for our research and with whom we co-published the findings from studies conducted in their schools (e.g., Sapienza et al., 2006). Throughout our work together, Dick challenged the research team during our regular Friday afternoon research meetings. He encouraged us to look beyond the obvious, to ask novel questions, and to interpret data that we had collected from a variety of perspectives. He also reminded us that in American education, bilingualism, becoming bilingual, and the encouragement of innovative language education programs within the core curriculum of public education are often viewed as *problematic*, *difficult*, and *undesirable*. He asked us to ponder the lead question shaping our projects and commitment to research: *What is the likelihood that students in American schools (primary, secondary, or tertiary) will graduate with bilingual proficiency and cross-cultural competence as a matter of course in the foreseeable future?*

This poignant question was at once generative for our thinking and simultaneously vexing for those of us believing in a curriculum that included foreign language proficiency outcomes. By posing this question, Dick pushed us to interrogate boundaries separating languages, cultures, and foreign language programs, which for many seemed impossible to breach. This overarching question fueled our

research and resulted in numerous articles on elementary school foreign language programs, content-based middle school language programs (e.g., Pessoa et al., 2007), and valid and reliable assessments of foreign language proficiency in children across the years of schooling (e.g., Igarashi et al., 2002; Wudthayagorn et al., 2002). Based on the findings of these studies in two schools, Dick and I co-authored a book – *A Tale of Two Schools: Developing Sustainable Foreign Language Programs* – which presented the cumulative linguistic and cultural achievements of two programs we investigated over several years. Our writing also addressed the perplexing issue of how to sustain foreign language programs for children in elementary schools at a time when programs seemed to be disappearing at alarming rates (Donato & Tucker, 2010).

A statement made by Zhang and Miller in their introduction to this volume clearly exemplifies Dick’s approach to our longstanding research on innovative language programming. They observe that it was through Dick’s dedication to boundary crossing – by his unwavering commitment to decenter anything “mainstream” – that we were able to gain deep and comprehensive insights into education as well as approaches to improving educational policy and practice. For many educators during this time, teaching children in more than one language and envisioning a curriculum that included well-articulated and extended K-12 sequences of language instruction seemed like an impossibility. But it was not so for Dick.

On a personal note, I owe a tremendous debt of gratitude to Dick. Closely working with him over the years taught me a great deal as I moved through the ranks of academia, especially when I assumed the role of chair in the Department of Instruction and Learning (now the Department of Teaching, Learning, and Leading). Dick’s influences on my personal and professional development are far too myriad to share in this small space, as many who have worked with Dick know all too well. Despite the years of reading research articles and learning about research methodology, it was Dick Tucker who taught me what it meant to conduct collaborative research that mattered and made a difference in the lives of children. So too was it Dick who showed me how to communicate clearly the value of our research and the direction it provided for the future to various constituents in schools and funding agencies. With deftness, conviction, and good humor, he consistently modeled how inflexible and unsurmountable boundaries could be crossed with only some effort and serious imagination.

Dick also taught me what it meant to be a mentor to a new generation of graduate students who would ultimately assume positions in universities and continue the work they had begun in their respective graduate programs. With gratitude, I recall the mentoring advice that Dick once gave me which has stayed with me throughout my career. *Doctoral students always need to move ahead in their work – never backwards*, he said. *Even when faced with the need to revise, rethink, or reconsider an issue, they must continue to move forward.* Along these same lines, Dick also liked to remind me that a dissertation is not good if it is not finished, another cautionary insight that many of his former doctoral students certainly remember.

As a highly accomplished administrator and academic leader, Dick additionally provided advice on the humanistic role of the chair that still resonates with me today. As he would always say *Support your staff, faculty, and students, learn about their work and areas of specialization, celebrate their successes, and make their accomplishments publicly known*. As a testimony to his beliefs, all staff members in Dick's office were recognized for their contributions and dedicated service through a university-wide award. Dick formed supportive personal relationships with students and colleagues that extended beyond department business, such as hosting dinners with his wife Rae in their home. He published one-page 'brag sheets' that summarized the yearly achievements of students and faculty and provided useful data on the current state of the department. It is fair to say that Dick's leadership at the university and for the profession should serve as the model for all those in administrative and leadership positions.

This volume is a tribute to G. Richard Tucker and to the countless ways he has crossed boundaries into the uncharted waters of language learning and educational policy and practice. Across these chapters, which are written by his previous graduate students, Dick's enduring legacy is clear. Supported by Dick's perspectives and careful mentoring, the research in this volume covers a wide range of topics including language learning and development, teacher education and its effect on instructional processes, program innovation and evaluation, and policy and planning for supporting literacy development and unraveling the linguistic complexities and language ideologies of study abroad programs. I congratulate Zhang and Miller for organizing this tribute to Dick into a volume that will clearly make a significant contribution to applied linguistics. I also thank them for including me in this important project honoring G. Richard Tucker on the occasion of his 80th birthday. This volume makes visible the significant impact that Dick has had over the years on the field of applied linguistics and on all of us who have had the life-changing privilege of working with him.

Department of Teaching, Learning, and Leading
University of Pittsburgh
Pittsburgh, PA, USA
donato@pitt.edu

Richard Donato

References

- Chinen, K., Donato, R., Tucker, G.R., & Igarahsi, K. (2003). Looking across time: Documenting middle school Japanese FLES students' attitudes, literacy, and oral proficiency. *Learning Languages*, 8(2), 4–10.
- Dominguez, R., Tucker, G. R., & Donato, R. (2005). Documenting curricular reform: Innovative foreign language education in elementary school. In P. Bruthiaux, D. Atkinson, W. G. Eggington, W. Grabe, & V. Ramanathan (Eds.), *Directions in applied linguistics: Essays in honor of Robert B. Kaplan* (pp. 56–71). Clevedon, UK: Multilingual Matters.
- Donato, R., Antonek, J. L., & Tucker, G. R. (1996). Documenting a Japanese FLES program: Ambiance and achievement. *Language Learning*, 46(3), 497–528.

- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable foreign language programs*. Clevedon, UK: Multilingual Matters.
- Donato, R., Tucker, G. R., & Hendry, A. H. (2015). Developing professional identities in applied linguistics: From doctoral study to professional practice. In Y. L. Cheung, S. B. Said, & K. Park (Eds.), *Advances and current trends in language teacher identity research* (pp. 217–234). London, UK: Routledge.
- Igarashi, K., Wudthayagorn, J., Donato, R., & Tucker, G. R. (2002). What does a novice look like: Describing the grammar and discourse of young learners of Japanese. *The Canadian Modern Language Review*, 58(4), 526–554.
- Pessoa, S., Hendry, H., Donato, R., Tucker, G. R., & Lee, H. (2007). Content-based instruction in the foreign language classroom: A discourse perspective. *Foreign Language Annals*, 40(1), 102–121.
- Sapienza, B. A., Donato, R., & Tucker, G. R. (2006). Learning a “second” foreign language: A district-wide foreign language program reaches the middle school. *The Language Educator*, 1(5), 24–27.
- Tucker, G. R., & Donato, R. (2003). Implementing a district-wide foreign language program: A case study of acquisition planning and curricular innovation. In D. Tannen & J. E. Alatis (Eds.), *Georgetown University Roundtable on Languages and Linguistics: 2001. Linguistics, language, and the real world: Discourse and beyond* (pp. 178–193). Washington, DC: Georgetown University Press.
- Tucker, G. R., Donato, R., & Antonek, J. L. (1996). Documenting growth in a Japanese FLES program. *Foreign Language Annals*, 29(4), 539–550.
- Tucker, G. R., Donato, R., & Murday, K. (2001). The genesis of a district-wide FLES program: A collaborative achievement. In R. L. Cooper, E. Shohamy, & J. Walters (Eds.), *New perspectives and issues in educational language policy: In honor of Bernard Dov Spolsky* (pp. 235–259). Philadelphia, PA: John Benjamins.
- Wudthayagorn, J., Donato, R., Tucker, G. R., & Igarashi, K. (2002). Self-assessment and the early language learner. *The Journal of Research Methodology*, 15(3), 387–409.

Foreword

When I received the invitation from co-editors Dongbo Zhang and Ryan Miller to write a foreword for this special volume in honor of G. Richard Tucker, I felt extremely grateful, humbled, and truly delighted have this opportunity. In addition to enjoying this occasion to share my personal appreciation for Dick's contributions, I know I am channeling the thanks of a very special group of students, mentees, and co-authors who have contributed to this volume, and the gratitude of many, many others, not only for the very visible role Dick has played, but also for Dick's less-than-visible and highly significant involvement in our personal and professional development. Greatly admired, respected, and beloved by countless colleagues, staff, and students who consider him their teacher, wise mentor, role model, and friend, Dick is the consummate scholar, educator, mentor, leader, and statesman with a huge, generous heart. Although there are many words, still inadequate, to describe Dick and his extraordinary impact, a place to start might be the title of this volume.

While certainly appropriate, the title *Crossing Boundaries in Researching, Understanding, and Improving Language Education: Essays in Honor of G. Richard Tucker*, invites attention to language choice and connotation. In this volume's introduction, Zhang and Miller present a well-constructed framework for considering the importance of boundary crossing "in researching and understanding the complexity of language education and improving language education policy and practice." Crossing boundaries can generate negative results or positive ones, depending on the context. Crossing boundaries can suggest inappropriate, unprofessional, or unethical behaviors, or perhaps uneventful, unremarkable movement, or, ideally, considerate, creative, path-opening progress. In every way, the definition for Dick and his life and work upholds the very best senses of crossing boundaries. Having the privilege to work with Dick has given many of us a glimpse into his playbook for crossing boundaries, and I venture to link his approaches with his love of family and love of football, too. Connecting with Dick professionally has meant for many of us seeing his love for his wife Rae and his family, and how he has modeled kindness and empathy for others who are navigating ups and downs and the boundaries of work and life outside of work. A former player and enthusiastic sports fan, Dick

is also the motivating coach who guides and cheers avidly for his teams, for his students, and for his colleagues to cross boundaries and score goals on the field, in the classroom, in research collaborations, and in administrative settings. He provides encouragement through challenges, and he celebrates with great joy the accomplishments of others. With respect for others' energies and constructive efforts, Dick has built solid bridges with his boundary crossings. Everywhere Dick has contributed, he has been a vital force promoting positive growth and a strong sense of community.

During the time he was paid (as he often put it) "for doing the work he loves," Dick spent his distinguished career on the faculty at McGill University, at the Ford Foundation as a Project Specialist in language centers around the world, then as President of the Center for Applied Linguistics for 13 years; and then he moved to Pittsburgh—the City of Bridges, how fitting!—to Carnegie Mellon, to become a faculty member, serving 12 years as Head of the Department of Modern Languages. After stepping out of the department head position, he "slowed down" to serve as Interim Dean of Students, Associate Vice Provost for Education—Qatar, Interim Dean of Carnegie Mellon Qatar, and Title IX Coordinator of the University. At Carnegie Mellon, while holding professional service and leadership positions in the wider profession, he also said "yes" to serve as chair of search committees for Dean of Students, Police Chief, Athletic Director, and Dean of Dietrich College, and all along, throughout his career, produced more than 200 scholarly publications; earned prestigious, well-deserved awards for his many accomplishments; and promoted the successful careers of hundreds of students and mentees. In retirement Dick has continued to engage actively and still receives frequent requests for his advice and support, and many of us continue to benefit from his generosity.

Yes, Dick is a well-known model of clarity and punctuality. In the days when there were more in-person meetings, I lost count of the number of times I know that individuals received follow-up thank you messages, directives, summaries, and confirmations from Dick by the time they had returned to their offices after meeting with him. With this remarkable efficiency Dick combines intelligent counsel, extraordinary respect for others, and genuine interest in their well-being. These qualities define him not only as a successful administrator, but also as a wonderful researcher and teacher.

In the Department of Modern Languages at Carnegie Mellon, where I have worked most closely with Dick, he has offered a unique perspective for crossing boundaries and bringing together diverse parts. He has enriched our unit with his broad perspective on language and language education. Unlike many universities across the country where there are separate departments of languages, our department houses programs in Arabic Studies, Chinese Studies, French and Francophone Studies, German Studies, Hispanic Studies, Japanese Studies, and Russian Studies, as well as two MA programs in Applied Second Language Acquisition and Global Communication and Translation, and a PhD program in Second Language Acquisition. Dick has displayed an understanding for the needs of each language area group and the undergraduate and graduate programs, and he has played a crucial role in building a very coherent unit and a congenial place with much

productive exchange and collaboration among students and faculty with research specializations in second language acquisition, literary and cultural studies, and technology-enhanced learning.

Dick takes seriously the mission of education at every level. With great dedication to his own students, both undergraduate and graduate, the content of his courses emphasizes issues related to education and learning processes, but what is noteworthy is how he has handled and shaped this content in the connections he has built with individuals and groups of individuals. One salient example is the undergraduate senior seminar that Dick designed and taught regularly in spring semesters. In this capstone course, Modern Languages majors from across language areas come together in their final semester to reflect on a variety of topics related to the experience of language learning. Students regularly commented to me as department head on their satisfaction with this course. They mentioned particularly Dick's extensive, deep knowledge and his superior organizational abilities as a teacher. From my conversations with Dick about students and teaching, it was clear that he established strong connections with his students, promoted their development, and often followed it and continued to be their advocate after their association in his courses. My own area of research focuses on Hispanic literary and cultural studies, and when I had the opportunity to teach the capstone seminar, I shaped my approach based on Dick's syllabus, and saw how wonderfully adaptable his course model was, even if one's area of specialization was not second language acquisition. When I proposed that we co-author an article about the course, he readily accepted, and the result of our collaboration was "Modern Languages Majors in the 21st Century: Broadening Disciplinary Frames of Reference and Global Awareness" (Polansky & Tucker, 2018). The capstone model champions crossing boundaries, and it allows for faculty members to shape the seminar in ways that can help the students see the importance of their courses of study, the significance of interdisciplinary competence, and their readiness for participation in the twenty-first century as well-prepared global citizens. Sharing updates about our students, Dick has expressed great delight in how creatively and productively our graduates have combined their study of modern languages with other disciplines, and how these choices have impacted their life paths.

Dick's view of teaching and education as a process of sharing and synthesizing is evident in his many collaborations beyond his own classroom. At Carnegie Mellon, he has been a dedicated advocate for building diversity and equity and addressing the needs of all students. In the Pittsburgh area, he was a key player in the union of a variety of constituencies to build a unique Spanish program in the Chartiers Valley School District. His impressive list of publications evidences a spectacular number of cooperative efforts, and clearly attests to his national and international recognition as an educator of educators.

Dick's love of data and extracting its key points connects interestingly with his fame for offering concise and often enumerated contributions for planning and implementing procedures and projects. A couple of his one-liners related to research and teaching of course have required more complex follow-up, but they distill foundational advice that, in practice, has proved crucial and forward-thinking. Especially

his graduate students, so many of them now successful researchers, teachers, and administrators, without sacrificing quality, and without losing sight of the ultimate significance of education and the interconnectedness of research and teaching, have lived by his often-repeated lines “A good dissertation is a finished dissertation” and “Put students first.” In meetings, sometimes free-flowing as meetings can be, his colleagues have deeply appreciated his succinct summaries at strategic moments, with his opening words “three things,” “four things,” or maybe “five things,” that shape the previous discussion and outline next steps. From his wide-ranging experience, Dick’s “Twelve Thoughts about Administration” have had far-reaching impact, and they have helped orient many of us in leadership positions. Dick’s administrative thoughts reflect very clearly his boundary crossings and how he has bridged his activities as an administrator, a researcher, and a teacher in and outside of the university. It’s not uncommon for lists to be aspirational, and those of us that know Dick and have seen this list would affirm without question that Dick the applied linguist has put his words into effective real-life practice. Here they are:

Twelve Thoughts About Administration

Department, College, University

- Delegate when possible and appropriate
- If you give people *responsibility*, give them *authority*
- Provide feedback frequently
- When decisions are made, write them down (and make sure the Administrative Associate/Business Manager has a copy)
- Acknowledge the special things that people do
- Schedule a regular meeting with Administrative staff to review events of the past week, tasks ahead, and problems anticipated
- Try to *respond* to all requests or queries *within 24 hours* if possible—positively or not
- Set firm, but realistic, deadlines for requests for information from faculty on various matters; and then proceed when the deadline has passed—with or without the input
- Ensure that people have information about ongoing activities, tasks, etc. (They don’t necessarily want to participate, but inevitably they do want to feel that they’re informed.)
- Establish and nurture the broadest possible base of personal contacts across *all* levels of university administration (a personal contact/request usually yields quick results)
- Study and understand the budget

Research, Education, and Networking in the Profession

Generally the above plus:

- Read broadly and file for future reference (with appropriate cross references so linkages can be made)
- Participate regularly in the core meetings across the field to remain abreast of people, trends, and current activities
- Maintain links with former colleagues, employees, and committee co-members, both domestically and internationally

The articles in this volume speak volumes about the breadth and depth of Dick's impact world-wide in the areas of language learning and development; language teacher activity and teaching; language program innovation, implementation, and evaluation; and language education policy and planning. I have had the pleasure of knowing many of the contributors to this volume since their first semesters as students at Carnegie Mellon, and it is awe-inspiring to see how they have crossed boundaries developing their careers and scholarship across continents and oceans and in this loving tribute to Dick. With Dick's impetus, in their activities and scholarly work, they are showcasing a bright future for this field as they pay forward with impressive productivity the many lessons learned. In this year of his 80th birthday, what a joy it is to honor Dick and extend to him a world of thanks!

Department of Modern Languages
Carnegie Mellon University
Pittsburgh, PA, USA
sp3e@andrew.cmu.edu

Susan G. Polansky

Reference

Polansky, S., & Tucker, G. R. (2018). Modern languages majors in the 21st century: Broadening disciplinary frames of reference and global awareness. *ADFL Bulletin*, 45(1), 139–149.

Contents

Boundary Crossing in Researching, Understanding, and Improving Language Education: An Introduction and the Tuckerian Impact	1
Dongbo Zhang and Ryan T. Miller	
Part I Language Learning and Development	
“I Want the Next Experience”: Israeli Adult Native Bilinguals Tell the Story of Their Childhood Bilinguality	35
Deborah Dubiner	
Boundary Crossing from the Start: 55 Years of Second Language Grammatical Gender Research in Review	57
Daniel R. Walter	
Non-Expert Raters’ Scoring Behavior and Cognition in Assessing Pragmatic Production in L2 Chinese	79
Shuai Li, Xian Li, Yali Feng, and Ting Wen	
Early Home and Community Support in Later Chinese Heritage Language Literacy Development	103
Haomin Zhang, Xi Cheng, and Jiexin Lin	
Part II Teachers and Instructional Processes	
Crossing the Disciplines: State of TESOL Teacher Education Programs in US Universities	127
Megumi Hamada and Ryan T. Miller	
Computerized Mediation in the Instruction and Development of L2 Pragmatic Competence: A Dynamic Assessment Perspective	149
Tianyu Qin	

Writing Development of the Case Analysis Genre: The Importance of Feedback and Negotiated Construction in the Teaching Learning Cycle 169

Maria Pia Gomez-Laich, Silvia Pessoa, and Ahmar Mahboob

Boundary Crossing: Integrating Visual Arts into Teaching Chinese as a Foreign Language 189

Yan Liu

Part III Program Innovation, Implementation, and Evaluation

Student-Level Variables and Academic Achievement in a Mandarin Dual Language Immersion Program. 213

Chan Lü, Amy E. Pace, and Liu Liu

Evaluating the Modular Curriculum of Chinese Language in Singapore Primary Schools: Insights from Students and Teachers. 231

Dongbo Zhang, Shouhui Zhao, and Xiaoxi Sun

Impact of the COVID-19 Pandemic on American Students’ Willingness to Study Abroad 257

Feng Xiao and Kun Nie

Writing as a Design Art: Crossing Boundaries Between Disciplinarity and Rhetoricity in a University Business Program 271

Ryan T. Miller, Silvia Pessoa, and David Kaufer

Part IV Language-in-Education Planning and Policy

EFL Literacy Development in Ethnic and Language Minority Learners: Implications from Tertiary-Level EFL Teaching and Learning in Ethnic Minorities in China 301

Sihui (Echo) Ke

“In a Foreign Bubble” in China: Language Use Among International Students During China’s *Belt and Road* 321

Yi Wang and Wenhao Diao

When Transfer Transfers: Applying Cross-Linguistic Reading Transfer Theory to Language of Instruction Policies in Low- and Middle-Income Countries. 341

Pooja R. Nakamura

Contributors

Xi Cheng The Psycholinguistics Lab, School of Foreign Languages, East China Normal University, Shanghai, China

Wenhao Diao Department of East Asian Studies, University of Arizona, Tucson, AZ, USA

Richard Donato Linguistics (University of Delaware), Newark, DE, USA
Department of Teaching, Learning, and Leading, School of Education, University of Pittsburgh, Pittsburgh, PA, USA

Deborah Dubiner Oranim College of Education, Kiryat Tiv'on, Israel

Yali Feng Department of Applied Linguistics/ESL, Georgia State University, Atlanta, GA, USA

María Pía Gómez-Laich Department of English, Carnegie Mellon University Qatar, Doha, Qatar

Megumi Hamada Department of English, Ball State University, Muncie, IN, USA

David Kaufer Department of English, Carnegie Mellon University, Pittsburgh, PA, USA

Sihui (Echo) Ke University of Kentucky, Lexington, KY, USA

Shuai Li Department of World Languages and Cultures, Georgia State University, Atlanta, GA, USA

Xian Li Department of Applied Linguistics/ESL, Georgia State University, Atlanta, GA, USA

Jiexin Lin The Psycholinguistics Lab, School of Foreign Languages, East China Normal University, Shanghai, China

Liu Liu College of Education, University of Washington, Seattle, WA, USA

Yan Liu Department of Asian and Middle Eastern Studies, Duke University, Durham, NC, USA

Chan Lü Department of Asian Languages and Literature, University of Washington, Seattle, WA, USA

Ahmar Mahboob Department of Linguistics, The University of Sydney, Sydney, Australia

Ryan T. Miller Department of English, Kent State University, Kent, OH, USA
Department of English, Ball State University, Muncie, IN, USA

Pooja R. Nakamura American Institutes for Research, Arlington, VA, USA

Kun Nie Department of Asian Languages & Literatures, Pomona College, Claremont, CA, USA

Amy E. Pace Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA

Silvia Pessoa Department of English, Carnegie Mellon University Qatar, Doha, Qatar

Susan G. Polansky Boston College, Boston, MA, USA
Department of Modern Languages, Carnegie Mellon University, Pittsburgh, PA, USA

Tianyu Qin Department of Modern & Classical Languages, University of North Georgia, Dahlonega, GA, USA

Xiaoxi Sun Department of Languages, Cultures and Visual Studies, University of Exeter, Exeter, UK

Daniel R. Walter Humanities Division, Emory University, Oxford College, Oxford, GA, USA

Yi Wang Department of Asian & Asian American Studies, Stony Brook University, Stony Brook, NY, USA

Ting Wen College of Chinese Studies, Beijing Language and Culture University, Beijing, China

Feng Xiao Department of Asian Languages & Literatures, Pomona College, Claremont, CA, USA

Dongbo Zhang School of Education, University of Exeter, Exeter, UK

Haomin Zhang The Psycholinguistics Lab, School of Foreign Languages, East China Normal University, Shanghai, China

Shouhui Zhao Department of Foreign Languages, University of Bergen, Bergen, Norway

Boundary Crossing in Researching, Understanding, and Improving Language Education: An Introduction and the Tuckerian Impact



Dongbo Zhang  and Ryan T. Miller 

We must describe the needs of the children and of the adults who desperately seek access to educational, social, and economic opportunities, not in the arcane and jumbled jargons so characteristic of academia, but with the precision, the elegance, and the simplicity of an artist. Only then can we hope to reach and to affect those responsible for the formulation of public policy. (Tucker, 2000c, p. 26)

Abstract This chapter first discusses the concept of boundary crossing and its learning potential in education and underscores the urgency of crossing a multitude of boundaries for researching, understanding, and improving language education. It then discusses the important role of expert boundary crossers. In particular, it highlights how G. Richard Tucker, whom this volume honors, has exemplified boundary crossing through his distinguished career in applied linguistics and language education for over half a century. The chapter ends with an overview of the four parts that form this volume and brief descriptions of how the chapters in the rest of the volume, each and collectively, contribute to language education research, policy, and practice through boundary crossing.

Keywords Boundary crossing · Language education · G. Richard Tucker

D. Zhang (✉)
School of Education, University of Exeter, Exeter, UK
e-mail: d.zhang4@exeter.ac.uk

R. T. Miller
Department of English, Kent State University, Kent, OH, USA
e-mail: rmill129@kent.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_1

1 Boundaries, Boundary Crossing, and Learning

The world is full of boundaries. There are visible geographical boundaries like physical borders between countries as well as less visible or invisible boundaries like socio-cultural differences that define communities. This is similarly the case for our inquiry into and interaction with the world around us. In educational sciences, there are diverse boundaries that often define who we are, what we do, and how we do things. We form distinct disciplinary communities (e.g., anthropology and sociology of education, policy studies, educational psychology, learning sciences, and cognitive neuroscience of education) and accordingly define our scholarly identities. There are also paradigmatic labels (e.g., positivism, post-positivism, and constructivism) and methodological approaches that define and characterize how we study, approach, and interpret educational issues. Educational research is further divided, and hence boundaries created, in accordance with diverse units of inquiry (e.g., from system and policy to school and classroom, and from the teacher to the student/learner), environments of education and modes of educational delivery, so on and so forth. Boundaries have been created, and continue to be created, voluntarily or involuntarily, which compartmentalize educational researchers and educators alike in defined zones. We are often confined by these boundaries socially, institutionally, and academically. These boundaries create discontinuities between theory and practice or between knowledge generation and sharing, as well as between various socio-cultural or educational settings, among many others. Such boundaries hinder the development of insights into the complexity of education and efforts to improve education.

Luckily, these boundaries can be crossed. Boundary crossing, as defined by Suchman (1994), refers to how professionals “enter into territory in which we are unfamiliar and, to some significant extent therefore unqualified” (p. 25). Boundary crossing restores continuity and brings learning potential (Ackerman & Bakker, 2011). Crossing boundaries involves constant reflection on and negotiation and contestation of ideas. Through boundary crossing, accepted ideas are scrutinized and challenged and new ones generated. Educational research has been influenced by anthropology, sociology, psychology, information and communication technologies, critical theories, policy studies, and more recently, data sciences, to name just a few. Interdisciplinary, collaborative work is being accepted as a new norm. The “paradigm war” has also been mitigated through the movement of mixed methods in educational research (Johnson & Onwuegbuzie, 2004). Through decentering anything “mainstream” and crossing boundaries, we gain more comprehensive and deeper insights into education as well as approaches to improving education policy and practice.

Ackerman and Bakker (2011), based on an integrative review of major studies on boundary crossing, concluded that four major mechanisms constitute the learning potential of boundary crossing, namely, identification, coordination, reflection, and transformation. Yet, crossing boundaries to achieve the learning potential is both challenging and risky. Stepping into an uncomfortable yet important zone requires

considerable renegotiation and reorientation. Crossing boundaries for expansive learning and an understanding of all sides of the education prism, particularly in a world that is quickly changing, is not easy. In boundary crossing, we “face the challenge of negotiating and combining ingredients from different contexts to achieve hybrid situations” (Engeström et al., 1995, p. 319). What boundaries to cross and how to cross them is both collective and personal. Navigating and negotiating these issues can be particularly daunting for junior scholars, who are often faced with a multitude of challenges and a multiplicity of positionings: personal, institutional, and academic. Expansive learning through boundary crossing, in a similar vein, applies to educational practices as well. Most if not all educational researchers are educators themselves, who are often faced with crossing boundaries of pedagogical ideas, educational systems, and institutional policies. To achieve the learning potential of boundary crossing and grow as a scholar and educator, or in Lave and Wenger’s (1991) words, to move from “legitimate peripheral participation” to a core member of “communities of practice,” the role of mentors and expert boundary crossers cannot be overstated (Barnett, 2008).

2 Boundaries in Language Education, G. Richard Tucker, and Boundary Crossing

The aforementioned outline of boundaries, boundary crossing, and learning underpins this volume. It has never been more important to cross boundaries in researching and understanding the complexity of language education and improving language education policy and practice. There are many “traditional” labels which we use, or boundaries created, to define who we are and what we do as language educators and/or researchers of language education. For example, in our research and practice, we differentiate between second, foreign, and heritage language; language majority vs. minority students; TESOL vs. World Languages (or Modern Foreign Languages); and within World Languages, “commonly taught languages” vs. “less-commonly-taught languages.” Programmatically, we differentiate between traditional foreign language programs, content-based instruction, and language immersion; and between programs for young school learners vs. university-based programs. Contextually, there is foreign language learning in a traditionally monolingual context vs. learning a language in a societal context or bilingual/multilingual societies. Disciplinarily, language education scholars also work with boundaries that define subfields: second language acquisition, classroom pedagogy, language policy and planning, language teacher education, language assessment and testing, to name just a few. Within the domain of language, there are further linguistic knowledge and skill labels that define what we do as researchers, not to mention the diverse languages we research and teach. There are also associations that seem to define further boundaries in our academic and professional life, such as NABE (National Association of Bilingual Education) vs. TESOL (Teaching English to

Speakers of Other Languages) International Association vs. ACTFL (American Council on the Teaching of Foreign Languages) vs AAAL (American Association for Applied Linguistics) vs. AERA (American Educational Research Association). Likewise, the “paradigm war” mentioned earlier has been specifically the case in language education research. While some scholars underscore researching and understanding cognitive processes of acquisition/learning of different linguistic skills in different learners, in different contexts, and through different mechanisms, others appreciate the social and semiotic nature of language learning and probe into the socio-cultural and political dimensions of language education (e.g., Kramsch, 2008; Pennycook, 2001; Van Lier, 2004).

Other than the labeling and boundaries above that may echo resonantly in our mind, there are also issues emerging in this quickly changing world that call for (re-)examination of boundaries that might have been taken for granted. For example, the COVID-19 pandemic, which broke out while this book was being prepared, has called our attention to the boundary between traditional, face-to-face teaching and learning of languages and technology-supported virtual learning. At a more hidden but deeper level, the values of intolerance of racism and xenophobia are at risk of being devalued through “othering” with the changing economic and political situations in the world. The anti-Asian racism following the initial outbreak of COVID-19 in China and other Asian countries, and the killing of George Floyd in the United States and the global resurgence of the Black Lives Matter movement in the midst of the pandemic are calling for further action on systemic racism. All these have (re)oriented us to critical issues around language, culture, society, and education, and to the importance of ambitious thinking and innovative practices in language education for promoting diversity, inclusion, equity, and social justice.

To research and understand the broad meaning (cognitive, social, cultural, humanistic, and political) of language education and improve policy and practice, the boundaries outlined above, and many others, must be crossed. The good news is that as a community, if not each individually, we have begun to cross some boundaries. For example, we have begun to research language learning as a complex dynamic system (Larsen-Freeman, 2012) and recognize that language education should be understood from an ecological perspective (Kramsch, 2008). SLA is being argued as a theory of practice (Hall, 1997). Efforts have also been taken to integrate approaches to language learning by bridging or crossing the boundaries between the so-called cognitive and sociocultural approaches (Atkinson, 2002; Douglas Fir Group, 2016; Hulstijn et al., 2014; Zuengler & Miller, 2006) and to deconstruct the traditional distinction between L1 and L2 for a “bilingual turn in SLA” (Ortega, 2010). Methodologically, it has also begun to be realized that the traditional boundaries between approaches – quantitative and qualitative – could and should also be crossed (King & Mackey, 2016), and mixed-methods research has begun to be underscored to bring new insights into the complexity of language learning and education (or applied linguistics) (Riazi & Candlin, 2014). The post-methods movement for language teaching has also been emphasized for researching, understanding, and innovating language pedagogies (Kumaravadivelu, 2003) by decentering any particular type of practice. Translingualism and translanguaging

have also contributed to our understanding about the dynamic, fluid, and hybrid nature of language learning and use and have brought new perspectives for understanding and interpreting issues of language learning and education and the role of language in education (García & Li, 2014).

Crossing the myriad of boundaries, and decentering whichever zones that define us or are defined by us, as mentioned earlier, is challenging and can be risky (see Donato et al., 2014; Tucker, 2000b). Luckily, there are brave, successful, and expert boundary crossers that have reminded us of the importance of boundary crossing (e.g., Tucker, 2000a, 2000b), exemplified boundary crossing to us, and inspired and mentored us to cross multiplicity of boundaries. G. Richard (Dick) Tucker, Paul Mellon University Professor Emeritus of Applied Linguistics at Carnegie Mellon University, is a notable one of them (Donato, 2013). Few scholars have valued and exemplified boundary crossing and demonstrated extraordinary success in it in language education research, policy, and practice more than Dick, whose career of over half a century has contributed to shaping the field of what we now know as applied linguistics (see Tucker, 2000b). Dick's over 200 publications (see [Appendix](#) for a selected bibliography) and many other types of scholarly contributions have taught us the importance of "a language competent society," and exemplified to us the importance and possibility of boundary crossing for researching and understanding language learning and education and improving policy and practice.

Dick started his academic career in the 1960s. The earliest and most visible boundary crossing to begin that journey was perhaps his decision to leave the United States to do his MA and PhD in psychology at McGill University, Canada, which according to him was made because "I could play intercollegiate football there." At McGill University, where he was later a member of the psychology faculty and a Professor of Psychology and Linguistics, Dick, in collaboration with colleagues and graduate students, crossed many linguistic, disciplinary, and methodological boundaries and published a number of studies that laid the foundation for understanding language learning processes, bilingualism, and program innovation and language teaching. The most notable boundary crossing during his tenure at McGill University was the collaborative work he conducted with Wallace E. Lambert on what was later known as the St. Lambert Experiment, a 12-year longitudinal evaluation of the effectiveness of French immersion programs in Quebec, Canada (Lambert & Tucker, 1972). While the St. Lambert project exemplified many dimensions of boundary crossing in language education, the most salient was perhaps between *research* on bilingual learning and innovative educational *practice* (program development, implementation, evaluation, and bilingual education policy, particularly language immersion or dual-language instruction) in schools.

This Tuckerian boundary crossing was later carried on to the Foreign Language in Elementary Schools (FLES) project that Dick conducted with his University of Pittsburgh collaborator Richard (Rick) Donato, who also wrote a foreword for this volume. In that project that spanned over a decade, Dick and Rick, in collaboration with their graduate students (e.g., Chinen et al., 2003; Donato & Tucker, 2010; Igarashi et al., 2002; Mitsui et al., 2007; Tucker et al., 1996), implemented and assessed an ambitious, multi-year, articulated program that taught Japanese and

Spanish to elementary school students in Pittsburgh, Pennsylvania. According to Donato (2013), Dick's "collaborative research in this area is the first to investigate empirically foreign-language learning in US elementary schools," and his studies "remain a primary source of information for school districts implementing foreign-language instruction across the grade levels" (p. 2).

The aforementioned boundary crossing between research and (innovative) practice is arguably only a fraction of the success that Dick has achieved. As evident in numerous "boundary objects" (projects, publications, presentations, policy engagements, leadership roles, etc.; Ackerman & Bakker, 2011), Dick, together with his collaborators, crossed a wide range of disciplinary, methodological, linguistic, institutional, programmatic, and national boundaries. The insights generated and discussed in his over 200 scholarly publications have convinced us that language learning and education are complex systems that necessitate synergistic insights into operations across diverse levels, ranging from policy, system, and standards to program, curriculum, and pedagogy, and from schools, administrators, and teachers to communities, parents, and students (see the introduction of each subsequent part of this volume for further detail).

To highlight, Dick's research has been informed by and contributed to a wide range of areas of scholarship that underpin language education, including, but not limited to, sociolinguistics, psycholinguistics, language acquisition, language policy and planning, bilingualism and multilingualism, language assessment and program evaluation, and teacher training and education. Dick's research has also crossed many boundaries between paradigms and methodological approaches to understanding, interpreting, and debating language learning processes (cognitive, social, and affective), language teaching, and policies: quantitative and qualitative; experimental and observational; historical and discursive; attitudinal survey and skill testing; to name just a few. In the FLES project (Donato & Tucker, 2010), for example, students' Japanese and Spanish proficiency was measured and monitored across many years of the programs to generate evidence of language development or effects of the program and instruction. In the meantime, classroom teacher-student talk was analyzed through discourse analysis, and the perspectives of multiple stakeholders – students, parents, and teachers – were elicited through different methods (e.g., interviews and questionnaires) to generate insights into sustainable implementation of early foreign language programs and policy implications for a language-competent society.

Dick's scholarship has also crossed boundaries of contexts (e.g., institutional, national, sociocultural, political) and programs (e.g., traditional TESOL and foreign/World Language programs, foreign language immersion, and content-based instruction). During his tenure at McGill University, Dick concurrently served as a language specialist for the Ford Foundation and conducted a number of studies on language use, policy, learning, and language teacher education in many societal and educational systems, from Southeast Asia and the Middle East to North Africa. During his directorship of the Center for Applied Linguistics (CAL), Dick adopted a "broad worldview," over and beyond what underpinned his earlier scholarship on French-English bilingual education in Canada, to unravel "the emergent

complexities of social policy and its relationship to second language acquisition in instructed contexts” (Donato, 2013, p. 2). He was also one of the four specialists who were sponsored by the US International Communication Agency and visited China to survey English teaching and teacher training there (Cowan et al., 1979). Dick contested the English-only movement in the United States and argued tirelessly for nurturing a language-competent society (e.g., Tucker, 1991, 1997). And together with his CAL colleague Jodi Crandall, Dick published a number of articles on language and content integrated instruction for language minority and language majority students in schools and universities (e.g., Crandall & Tucker, 1990; Tucker & Crandall, 1989).

Crossing linguistic boundaries also saliently characterizes Dick’s scholarship and engagement in language education policy and practice. In addition to using bilingual/multilingual lens to unravel the complexity of language processing, learning, and education (e.g., Bruck et al., 1974; Tucker, 1998, 2001), Dick, together with his collaborators and graduate students, investigated the learning and teaching of diverse languages across national, institutional, and programmatic contexts and types of learners (e.g., school children vs. university students; language minority vs. majority students; heritage vs. foreign language learners). Those languages include Arabic, Chinese, English, Filipino, French, Hebrew, Japanese, and Spanish, to name just a few. During his tenure at Carnegie Mellon University (CMU), where he served multiple academic and leadership roles before retirement in 2015, Dick also aimed to bridge language and global education for university students. He and his CMU colleagues outlined global literacy and argued for the cultivation of it in American university students in response to the multitude of challenges with which the world is faced today (Nair et al., 2012; Polansky & Tucker, 2018).

In addition to his distinguished research and scholarship, Dick also exemplifies the role of a boundary crosser in promoting communication between diverse stakeholders and communities for understanding the critical importance of language learning, bilinguality, and education. He crossed institutional boundaries (from policy-informing institutes to academic associations and from higher education to schools) and the boundaries between the many academic, educational, administrative, and leadership roles he served. While working at McGill University (1968–1978), Dick, as noted earlier, was concurrently a language education advisor for the Ford Foundation and conducted a number of studies on language use, policy, and education in Southeast Asia, the Middle East, and North Africa. After he left McGill University, Dick became the Director of CAL in Washington DC (1978–1992). During his tenure as Director, CAL established the National Network for Early Language Learning to promote foreign language instruction in elementary schools in the US. Under his leadership, the scope of the center was significantly expanded, the annual budget significantly increased, and staff nearly tripled. CAL later established the Tucker Fellowship in 1992 in honor of his distinguished service and leadership.

Dick joined Carnegie Mellon University in 1992 as a Professor of Applied Linguistics and became in 1995 the Head of the Department of Modern Languages, which now offers eight languages to CMU students (Arabic, Chinese, French,

German, Italian, Japanese, Russian, and Spanish), 50% of whom (the national average being 9% per university), accordingly to a report of the Modern Language Association (MLA) in 2012,¹ took at least one foreign language class while studying at CMU. Dick served as the Head of the department for 12 years, after which the Headship passed on to Susan G. Polansky, Dick's long-term colleague and collaborator (see, for example, Polansky & Tucker, 2018) and also the author of a foreword for this volume. Since 2007, Dick continued to serve multiple leadership roles at CMU until he retired in 2015 as "Mr. Everything" and the Paul Mellon University Professor Emeritus of Applied Linguistics.² These roles included, for example, the Interim Dean of Student Affairs, Associate Vice Provost for Education for Carnegie Mellon Qatar, Interim Dean of Carnegie Mellon University Qatar, and Title IX Coordinator. Dick received the Elliott Dunlap Smith Award from the College of Humanities and Social Sciences for distinguished teaching and educational service (1999), and later won the University's Doherty Award for sustained contributions to excellence in education (2007).

Dick has also crossed boundaries between professional associations or organizations with multiple leadership roles that facilitated research and communication on language learning, language teaching and program innovation, and educational policy. Among many other notable roles, Dick was a member of the Board of Directors of TESOL International Association (2003–2006), and is a foundation trustee of the International Research Foundation for English Language Education (TIRF), a nonprofit organization which, according to its position statement, gives "high priority to the development of a coherent program of language learning research, teaching research, and information dissemination." Because of his distinguished scholarship, leadership, and service, Dick won prestigious awards or recognitions from all major language education associations in the United States, including the American Association for Applied Linguistics (Distinguished Scholarship and Service Award; 2003), the American Council on the Teaching of Foreign Languages (Paul Pimsleur Award for Distinguished Research; 1997), and TESOL International Association (The James Alatis Award for Service to TESOL; 1998). He was also chosen by the National Association for Bilingual Education (NABE) as the "Honoree of the Year" (1995) for his "significant contributions to the body of research on language acquisition and the establishment of sound bilingual education programs."

¹ <https://thetartan.org/2012/12/3/news/foreignlanguage>

² <https://www.cmu.edu/dietrich/modlang/news-stories/2015/tucker-retires.html>

3 Goals of this Volume and Contributors/Boundary Crossers

The goals of this volume are two-fold. Firstly, we aim to use original research papers from authors who are on the frontline of language education and research to explore, exemplify, and discuss boundary crossing, and through that boundary crossing to generate new insights that improve language education, policy and practice. Although there are several special issues of journals or volumes that shed light on paradigmatic hybridity for language education or applied linguistics research (e.g., Kostoulas, 2019), that focus is necessarily restricted with respect to the multiplicity of boundary crossing. Springer's multiple-volume *Encyclopedia of Language and Education*, for which Dick and David Corson (Tucker & Corson, 1997) edited the volume on second language education for the first edition, is perhaps the most ambitious project that shows the landscape of language (in) education and sheds light on boundaries for crossing. Yet, the entries in the volumes did not specifically intend to explore, exemplify, and discuss boundary crossing between languages, programs, contexts, learners, units of inquiry, etc. In this respect, the present volume fills a niche.

Secondly, we aim to honor Dick's distinguished scholarship on language education and pay tribute to his inspiration and mentorship that have encouraged and scaffolded our crossing of boundaries academically and professionally. Dick is an outstanding boundary crosser; a tireless advocate on what language learning and bilinguality mean to who we are as an individual, a community, and a society; an eminent scholar and professor; and an inspirer, role model, and selfless mentor to the contributors of this volume (and, needless to say, many others in the fields of language education and applied linguistics). The courage, ambition, and success of the path exemplified in Dick's career has inspired us, and will continue to inspire us, to cross boundaries to research and understand the complexity of language education and improve policy and practice.

The authors come from diverse backgrounds. They are from different places in the world (e.g., China, Israel, Qatar, UK, USA); they have taught diverse languages and speak and research even more, such as Arabic, Chinese, English, French, German, Hebrew, Japanese, Kannada, and Spanish, to name just a few; their contexts of research and practice – national, sociocultural, institutional, and programmatic – are also diverse. Their students and research participants also vary, ranging from language minority to language majority students and from linguistically and culturally diverse students to students in a traditionally monolingual setting, etc. Their research is informed by various theoretical perspectives and methodological approaches. Yet, they “cross boundaries” and come together for this volume because they are all, like Dick, boundary crossers in language education research and practice, and more importantly, because the lead authors all share the same path of receiving their doctoral degree from the PhD in Second Language Acquisition program at Carnegie Mellon University where they were taught, advised, inspired, and mentored (and continue to be mentored) by Dick. We view this volume as an outcome of our collaborative action research, under the mentorship of Dick, to explore

boundary crossing as language education researchers and language (teacher) educators.

4 Boundary Crossing Characterizing This Volume

Boundary crossing is characteristic in this volume of individual chapters and parts, and across these chapters and parts. For the *ad hoc* purpose of organizing and structuring chapters, we use four broad themes of language education to organize this volume into four parts, namely, language learning and development; teachers and instructional processes; program innovation, implementation, and evaluation; and language-in-education policy and planning. These parts or themes also cover the major areas of language education where Dick has made distinguished contributions. Specifically, boundary crossing is characterized in this volume in the following three ways (see Fig. 1).

First, while the four parts are separately presented, the themes are necessarily cross-cutting. In other words, boundaries are crossed between the areas of scholarship that often define our niche, scholarly identity. For example, research on teachers and teaching (Part II) may be contextualized in the implementation of an innovative program (Part III) or more broadly in shifting policies and may inform policy-related decisions (Part IV). Likewise, program evaluation (Part III) may well involve collecting evidence on student learning and language development (Part I) and classroom processes (Part II). This type of boundary crossing is clearly exemplified in Dick's scholarship, as discussed in detail in the introduction of each part of this volume.

Second, the chapters that form a part, despite a shared focus on the broad thematic issue, approach that issue by crossing boundaries of languages, methodologies, programmatic contexts, and socio-political or educational systems, among others. For example, while all chapters of Part IV focus on language-in-education planning and policy, they are informed by diverse theoretical perspectives and/or adopt different methodological approaches to understanding the interplay of many micro and macro factors in different social or educational settings.

Finally, and most importantly, each individual chapter manifests boundary crossing within the chapter itself. For example, to address the complexity of language learning processes and development, a chapter in Part I may bridge theoretical frameworks, adopt cross-linguistic perspectives and designs, and/or "mix" methods. The most distinctive feature of this volume is that all chapters explicitly address and discuss boundary crossing, which may be either foregrounded and directly frame a study or, in a less direct way, be encapsulated in the discussion of the study and its findings where boundaries crossed are discussed to highlight the insights generated into language education.

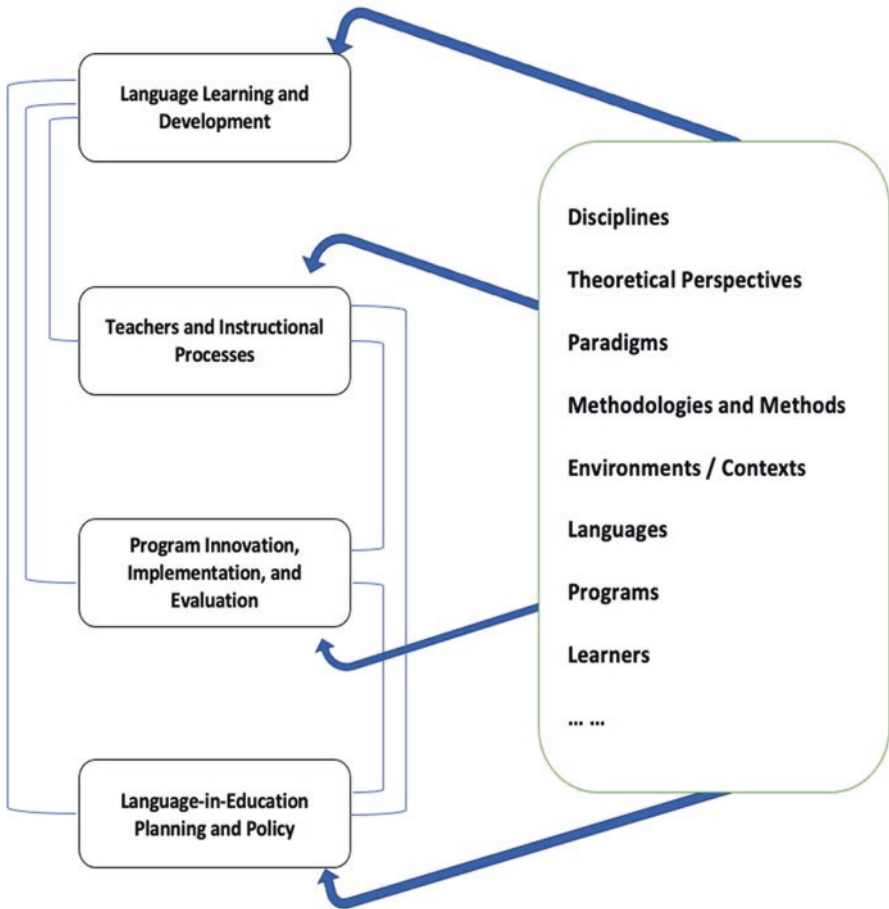


Fig. 1 Boundary crossing characterizing Dick Tucker’s scholarship and this volume

5 Volume Organization and Introduction to Chapters

This introduction chapter (Chap. 1) is preceded by the forewords of Dr. Richard Donato and Dr. Susan G. Polansky on their respective journey of collaboration and boundary crossing with Dick. Rick is a Professor and was Chair of the Department of Instruction and Learning (now the Department of Teaching, Learning, and Leading) in the School of Education at the University of Pittsburgh and Dick’s long-term research collaborator, notably in the FLES project described earlier. Susan had been Dick’s colleague and co-author at Carnegie Mellon University and succeeded Dick as the Head of the Department of Modern Languages, a role she served until 2020. This introduction chapter is followed by the four parts, each beginning with an introduction that outlines what we call the Tuckerian impact and followed by four chapters. These chapters cut across themes of the parts and address, from

different perspectives, boundary crossing in researching, understanding, and improving language education that emulates the Tuckerian impact.

The first part focuses on language learning and development. In the introduction, the boundary crossing exemplified in Dick's scholarship on language learning, learners, and bilinguality is presented, particularly his research that transcended boundaries of languages and language programs and the paradigmatic pluralism that he underscored for understanding the cognitive and social underpinnings of becoming bilingual and the educational implications of bilingualism and bilinguality.

In Chap. 2, Dubiner explores the intersectionality of language/bilinguality and ethnic and national identity, with a focus on narratives of adult Israelis who were born to immigrant parents right around Independence. The lived experience of those participants from different social and familial backgrounds showed how they crossed boundaries of home language and Hebrew in identity construction (diasporic vs. Israeli identity) during the country's revival of Hebrew as the national language.

In Chap. 3, Walter crosses linguistic, theoretical, and methodological boundaries, among others, to review grammatical gender across languages and the research on the acquisition of grammatical gender in L2 learners. The author also deconstructs the boundary between SLA theory and instruction and discusses how boundaries between pedagogies such as functional and sociocultural approaches can be crossed for the teaching of grammatical gender.

Chapter 4 addresses the interface between language assessment, acquisition, and use with a focus on L2 Chinese learners' pragmatic production. Li and colleagues contend that L2 pragmatic assessment often relies on expert raters and focuses predominantly on speech acts. This has constrained understandings about learners' real-world language use where non-expert, native speakers are usually the interlocutors and "assessors." A study was thus conducted to cross boundaries by probing into non-expert raters' scoring behavior and cognition in assessing L2 learners' pragmatic production that included both speech acts and pragmatic routines.

In Chap. 5, Zhang and colleagues explore the complex interplay of factors that influence Chinese as a heritage language (HL) reading development. Their study found that HL learners' language and literacy experiences in the community had a more salient effect on their literacy development than those at home. It underscored the importance of crossing boundaries between learner-internal, resource factors and learner-external, socio-contextual factors for understanding HL literacy development and maintenance.

Each paper exemplifies overlapping yet distinct boundary crossing, which has contributed to our understanding about the complex process, and the meaning, of learning languages and becoming bilingual. Additionally, they cross boundaries of the thematic areas that form the parts of this volume. Chapters 3 and 4, for example, clearly inform language instruction and assessment (Part II). Likewise, Chap. 2 touches on language, identity and nation building in a broad policy context (Part IV); and the findings of Chap. 5 also shed light on HL maintenance with policy implications (Part IV). Collectively, these chapters also cross boundaries of languages and contexts, as well as the boundary between theory and practice.

The second part focuses on language teachers and teaching, including teacher learning and professional development. The introduction presents the contribution of Dick's scholarship and boundary crossing to our understanding about innovations in language teaching and language teacher education. Among the many boundaries crossed (e.g., programs and national contexts, learners, and methodological pluralism for researching language teaching and teachers), highlighted is Dick's emphasis that language teaching is both science and art, and language educators need to be both a scientist and an artist to create an impact.

Chapter 6 focuses on K-12 English learner (EL) teacher education programs in the United States. Hamada and Miller, based on an examination of programs' offerings, compare the knowledge bases that different university-based EL teacher education programs draw on in the United States. They demonstrate how factors such as home department (language vs. education) and program level (baccalaureate vs. post-baccalaureate / master's) influence whether an EL teacher education program is more applied linguistics- or general education-oriented. The authors argue that programs should cross disciplinary boundaries to integrate research and practices from different fields for educating EL teachers in US K-12 schools.

Chapter 7 focuses on transcending the boundary between the teacher as a human mediator in dynamic assessment (DA) of student learning and the mediational role of carefully designed digital tools based on intelligent computing. Qin explores computerized mediation through DA to facilitate the development of the ability to comprehend implied meaning in L2 learners of Chinese. She discusses how computerized mediation tools can cross boundaries of learning environments and traditional teacher-student roles for ubiquitous learning of languages.

In Chap. 8, Gómez-Laich and colleagues aim to cross disciplinary boundaries between subject learning and writing in English for Academic Purposes (EAP) in universities with English as the Medium of Instruction (EMI). They report a study in which writing faculty/applied linguists worked in collaboration with information systems (IS) faculty and adapted the Teaching Learning Cycle to scaffold the writing of the case analysis genre in two IS courses in an American university in the Middle East. The authors also discuss the strategies they have used for promoting that boundary crossing.

Chapter 9 situates boundary crossing in a university-based Chinese as a Foreign Language (CFL) program in the United States. Liu aims to integrate visual arts into her own CFL teaching and explores the relations between humanities, art appreciation and creativity, and language education. This chapter reports on the author's effort to cross disciplinary boundaries and boundaries of instructional environments in her CFL teaching and presents a mixed-methods study that evaluated that effort with a focus on student motivation and perceptions.

Each chapter in this part exemplifies boundary crossing that contributes to our understanding about language teaching, teachers, and teacher education. The chapters also cut across the themes of the four parts. For example, Chap. 6, while focusing on disciplinary influences on EL teacher education, has implications for teacher education programs and the evaluation of these programs (Part III). Likewise, while Chapters 8 and 9 have a clear focus on language teaching, they also explore

students' language learning and use (Part I) and generate insights into program innovation (Part III). Needless to say, the four chapters collectively also cross boundaries of languages, programs and contexts, and research methods.

The third part consists of four chapters that focus on language program innovation, implementation, and evaluation, an area where Dick has made a highly distinguished contribution. Like the previous parts, this part first outlines the boundary crossing exemplified in the Dick's scholarship and discusses the Tuckerian impact. A very characteristic if not the most characteristic type of boundary crossing in program implementation and evaluation is that it is a collective enterprise, as has been underscored by Dick, that involves engagement across multiple stakeholders, such as scholars / evaluators, teachers, administrators, communities, parents, and of course students / learners.

In Chap. 10, Lü and colleagues contextualize their boundary crossing in different Dual Language Immersion (DLI) programs in the United States. They argue that even though evidence has been established on benefits of DLI programs on students' academic achievement, most studies on these educational benefits focused on Spanish DLI programs. The authors address this niche by analyzing the longitudinal data they collected from a Mandarin immersion program in urban California that enrolled ethnically and linguistically diverse students. Lü et al. found that students' math and English language arts scores showed significant growth over time and the trajectory of growth was not influenced by students' language background and race/ethnicity.

Chapter 11 underscores the importance of methodological hybridity and insights of stakeholders in curriculum/program evaluation and evidence-based language education policies and planning. Zhang and colleagues contextualize the discussion of boundary crossing in program innovation and evaluation in Singapore, where school curriculums are constantly reviewed and reformed to meet with the realities of the evolving sociolinguistic landscape and where they conducted a multi-year project to evaluate the Chinese Modular Curriculum in primary schools. The authors report their findings on teachers and students and discuss the interface between sociolinguistics, language policy and planning, and curricular/program innovation and reform.

Chapter 12 does not involve evaluation of a program or a national curriculum but focuses on innovation and implementation of study abroad (SA) programs, which were heavily impacted by the COVID-19 pandemic. Xiao and Nie conducted a questionnaire survey of American college students, the most important stakeholder of SA programs, to understand their perceptions of SA in academic study, willingness to study abroad, and the impact of SA (or potentially the lack thereof) on their career prospects and life during a time full of uncertainties. Students' insights are discussed to help departments and SA program directors evaluate current curricular requirement and explore new, innovative models and practice.

In EMI universities, programs should aim for both disciplinary learning and professional communication skills in a discipline; yet, as Miller, Pessoa, and Kaufer point out in Chap. 13, boundaries often exist between *writing to learn* and *learning to write* views in program faculty. The authors argue for a writing as design approach

to disciplinary writing instruction and report how this approach was exemplified in a business content course in an EMI university in the Middle East. They interviewed instructors and analyzed instructional materials and student writing, and, based on the findings, they make recommendations on how business programs can leverage a writing as design conceptualization to improve student writing and bridge any boundaries perceived by faculty (and perhaps students as well) between disciplinarity and rhetoricity.

The four chapters in this part, while each individually contributing to program innovation, implementation, or evaluation through boundary crossing, also cross thematic boundaries of the four parts. Chapters 10 and 11, while showing a focus on program evaluation, both have strong implications for language education planning and policy (Part IV). Chapter 10 also informs language learning and bilinguality (Part I) and Chap. 11 also contributes to understandings about language teachers and teaching (Part II). Likewise, Chapters 12 and 13 both shed light on curriculum or program policies (Part IV). The four chapters of course also collectively cross boundaries of languages, programs and educational contexts, research methods, so on and so forth.

The fourth part explores boundary crossing in language education policy and planning. Dick's publications on language education policy and planning and engagement in policy formulation, discussion, and debate are a very salient contribution of his to language education and applied linguistics. This is discussed in the part introduction. In particular, for boundary crossing exemplified in Dick's scholarship, the introduction highlights his global view on language education, bi-/multilingualism, and policy and planning. It is emphasized that this global view, or a view that transcends sociolinguistic, national, and educational contexts, is fundamental for us to understand the many micro and macro factors that interplay in policy formulation, implementation, and contestation.

Chapter 14 focuses on EFL learning and literacy in ethnic minority (EM) students who learn English as a third language in China. Ke first conducts a systematic review of relevant studies to identify the challenges EM students encounter in university EFL learning and effective models/programs/pedagogies for EM students. She then reports a study that compared the English literacy profiles of EM and Chinese-speaking Han-majority university students. Based on the review and empirical findings, Ke argues that EFL instruction for EM students needs to consider their varied backgrounds and linguistic repertoires and avoid deficit-perspectives, and calls for collaborative, integrated approaches to EFL policy and practice for EM students in multilingual settings.

In Chap. 15, using the metaphor of a "foreign bubble," Wang and Diao characterize international students' creation of an English-speaking environment for academic purposes in an EMI program in Shanghai, China. This English-speaking "foreign bubble," which appears to justify the students' avoidance of Chinese, however, is contested by other program stakeholders. The authors engaged with students, faculty, as well as staff of the program to unravel this unique policy context of language in education, which is being promulgated in response to China's goal of

internationalizing its higher education through EMI education and in service to the country's global strategies such as the Belt and Road Initiative.

A distinct focus of Chap. 16 is its global perspectives on language education, bi-/multilingualism, language policy and planning. Nakamura discusses how theoretical concepts of cross-linguistic transfer in bilingual reading development could and should be used to address urgent issues of socioeconomic advancement and educational access for all in the developing world. The discussion is situated in the complex reality that in many low- and middle-income countries (LMICs), which are also typically multilingual, the disparity between home and instructional languages has constrained literacy development, academic achievement, and socio-economic mobility. With references to a range of multilingual countries, Nakamura discusses how "cross-linguistic reading transfer theory" is valuable for addressing issues of initial literacy instruction and transition to literacy instruction in other languages. The author also discusses the importance of stakeholder engagement and boundary crossing for applications of theory to policy and practice for improving literacy outcomes in low-resourced LMIC contexts.

Inquiries into language education planning and policy need to engage with policy contexts and stakeholders or actors. In this respect, the foregoing three chapters all cross the thematic boundaries of this volume, yet each provides insights into boundaries and boundary crossing in distinct ways, crossing boundaries of perspectives, languages, and methodological approaches. Chapter 14 shows a focus on language learning and proficiency in EM students (Part I) and Chap. 15 teachers/faculty and programs (Parts II and III). Chapter 16 shows a notable, global perspective. It cuts across the themes of all parts to show that policy and planning in language (in) education is a complex system that requires crossing boundaries between theories and practice; socioeconomic and sociopolitical dimensions of language education; program models and evidence-based innovation; and diverse policy actors.

6 Conclusion

Language education, like general education (Jacobson et al., 2019), is a complex system. Understanding the system and improving practice and policy require crossing diverse boundaries and decentering any particular disciplinary tradition, paradigm, or approach; language; learner; learning environment; programmatic, institutional, and political setting; so on and so forth. We hope this book has achieved its purposes of underscoring and exemplifying boundary crossing in language education research; honoring Dick as a distinguished scholar, expert boundary crosser, and mentor; and celebrating Dick's eightieth birthday and his commitment to language education research, policy, and practice for over half a century. We also hope that this book has shown the learning potential in boundary crossing for professional development of language education researchers and language (teacher) educators.

Appendix G. Richard Tucker: A Selected Bibliography

This appendix presents a selected bibliography of G. Richard Tucker's scholarly works that have been produced from when he received his master's degree in psychology at McGill University in 1965 to around when he retired in 2015 as the Paul Mellon University Professor of Applied Linguistics Emeritus at Carnegie Mellon University. These works are selected from over 200 scholarly works that Dick has produced. We hope that this bibliography supplements the descriptions that have been presented in this chapter to show how Dick has crossed boundaries in generating insights into language education, how his scholarship has contributed to understanding and improving language education, and how his works have contributed to shaping the field that we now define as applied linguistics.

1965–1967

- Anisfeld, M., & Tucker, G. R. (1967). English pluralization rules of six-year-old children. *Child Development*, 38(4), 1201–1217.
- Cohen, S. P., Tucker, G. R., & Lambert, W. E. (1967). The comparative skills of monolinguals and bilinguals in perceiving phoneme sequences. *Language and Speech*, 10(3), 159–168.
- Hayes, A. S., Lambert, W. E., & Tucker, G. R. (1967). Evaluation of foreign language teaching. *Foreign Language Annals*, 1(1), 22–44.
- Lambert, W. E., Frankle, H., & Tucker, G. R. (1966). Judging personality through speech: A French-Canadian example. *Journal of Communication*, 16(4), 305–321.
- Tucker, G. R. (1965). *The nature of English pluralization rules of kindergarten children*. Unpublished MA thesis. McGill University.
- Tucker, G. R. (1967). *French speakers' skill with grammatical gender: An example of rule-governed behavior*. Unpublished doctoral dissertation. McGill University.

1968

- Tucker, G. R. (1968). Judging personality from language usage: A Filipino example. *Philippine Sociological Review*, 16(1/2), 30–39.
- Tucker, G. R., Lambert, W. E., Rigault, A. (1968). A psychological investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, 7(2), 312–316.

1969

- Tucker, G. R., Lambert, W. E., Rigault, A. (1969). Students' acquisition of French gender distinctions: A pilot investigation. *IRAL - International Review of Applied Linguistics in Language Teaching*, 7(1), 51–55.

1970

- Tucker, G. R., & Gedalf, H. (1970). Bilinguals as linguistic mediators. *Psychonomic Science*, 20(6), 369–370.
- Tucker, G. R., & Lambert, W. E. (1970). The effect on foreign language teachers of leadership training in a foreign setting. *Foreign Language Annals*, 4(1), 68–83.

Tucker, G. R., Otones, F. T., & Sibayan, B. P. (1970). An alternate days approach to bilingual education. In J. A. Atlatis (Ed.), *Georgetown University round table on languages and linguistics 1970* (pp. 281–295). Georgetown University Press.

1971

Bruck, M., Jakimik, J., & Tucker, G. R. (1971). Are French immersion programs suitable for working-class children? A follow-up investigation. *Word*, 27(1–3), 311–341.

Capco, C. S., & Tucker, G. R. (1971). Word association data and the assessment of bilingual education programs. *TESOL Quarterly*, 5(4), 335–342.

d'Anglejan, A., & Tucker, G. R. (1971). Academic report: The St. Lambert Program of home-school language switch. *Modern Language Journal*, 55(2), 99–101.

Davine, M., Tucker, G. R., & Lambert, W. E. (1971). The perception of phoneme sequences by monolingual and bilingual elementary school children. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, 3(1), 72–76.

Gatbonton, E. C., & Tucker, G. R. (1971). Cultural orientation and the study of foreign literature. *TESOL Quarterly*, 5(2), 137–143.

Hebb, D. O., Lambert, W. E., & Tucker, G. R. (1971). Language, thought and experience. *Modern Language Journal*, 55(4), 212–222.

Tucker, G. R., & d'Anglejan, A. (1971). Some thoughts concerning bilingual education programs. *Modern Language Journal*, 55(8), 491–493.

Tucker, G. R., Taylor, D. M., & Reyes, E. (1971). Ethnic group interaction in a multiethnic society. *International Journal of Psychology*, 6(3), 217–222.

1972

Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Newbury House Publishers.

Seligman, C. R., Tucker, G. R., & Lambert, W. E. (1972). The effects of speech style and other attributes on teachers' attitudes toward pupils. *Language in Society*, 1(1), 131–142.

Tucker, G. R., & Lambert, W. E. (1972). Sociocultural aspects of foreign-language study. Northeast conference on the teaching of foreign languages. In J. W. Dodge (Ed.), *Other words, other worlds: Language-in-culture* (pp. 26–30). The Northeast Conference on the Teaching of Foreign Languages.

1973

d'Angelian, A., & Tucker, G. R. (1973). Communicating across cultures: An empirical investigation. *Journal of Cross-Cultural Psychology*, 4(1), 121–130.

d'Anglejan, A., & Tucker, G. R. (1973). Sociolinguistic correlates of speech style in Quebec. In R. W. Shuy & R. W. Fasold (Eds.), *Language attitudes: Current trends and prospects* (pp. 1–27). Georgetown University Press.

Rémillard, L., Tucker, G. R., & Bruck, M. (1973). The role of phonology and lexicon in eliciting reactions to dialect variation. *Anthropological Linguistics*, 5(9), 383–397.

Lambert, W. E., Tucker, G. R., & d'Anglejan, A. (1973). Cognitive and attitudinal consequences of bilingual schooling. *Journal of Educational Psychology*, 65(2), 141–159.

Tucker, G. R., Lambert, W. E., & D'Angelian, A. (1973). Are French immersion programmes for working class children? A pilot investigation. *Language Sciences*, 25, 19–26.

1974

Bruck, M., Lambert, W. E., & Tucker, G. R. (1974). Bilingual schooling through the elementary grades: The St. Lambert Project at grade 7. *Language Learning*, 24(2), 183–204.

Campbell, R. N., Taylor, D. M., & Tucker, G. R. (1974). Teachers' views of immersion-type bilingual programs: A Quebec example. *Foreign Language Annals*, 7(5), 106–110.

Saegert, J., Scott, S., Perkins, J., & Tucker, G. R. (1974). A note on the relationship between English proficiency, years of language study and medium of instruction. *Language Learning*, 24(1), 99–104.

Scott, M. S., & Tucker, G. R. (1974). Error analysis and English-language strategies of Arab students. *Language Learning*, 24(1), 69–97.

Stubbs, J. B., & Tucker, G. R. The cloze test as a measure of English proficiency. *Modern Language Journal*, 58(5/6), 239–241.

Tucker, G. R. (1974). *Innovative approaches to second-language teaching*. TESOL.

Tucker, G. R. (1974). Methods of second-language teaching. *Canadian Modern Language Review*, 31(2), 102–107.

1975

Bruck, M., Lambert, W. E., Tucker, G. R., & Bowen, J. D. (1975). The 1968 NDEA Philippine institute for TESL teachers: A follow-up evaluation. *Foreign Language Annals*, 8(2), 133–137.

d'Anglejan, A., & Tucker, G. R. (1975). The acquisition of complex English structures by adult learners. *Language Learning*, 25(2), 281–296.

El-Dash, L., & Tucker, G. R. (1975). Subjective reactions to various speech styles in Egypt. *International Journal of the Sociology of Language, Issue 6*, 33–54.

Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communication skills of bilingual children. *Child development*, 46(4), 1010–1014.

Harrison, W., Prator, C., & Tucker, G. R. (1975). *English-language policy survey of Jordan*. Arlington, VA: Center for Applied Linguistics.

Markman, B. R., Spilka, I. V., & Tucker, G. R. (1975). The use of elicited imitation in search of an interim French grammar. *Language Learning*, 25(1), 31–41.

Tucker, G. R. (1975). The acquisition of knowledge by children educated bilingually. In D. P. Dato (Ed.), *Georgetown University round table on languages and linguistics 1975* (pp. 267–277). Georgetown University Press.

Tucker, G. R., & d'Anglejan, A. (1975). New directions in second language teaching. In R. C. Troike & N. Modiano (Eds.), *Proceedings of the First Inter-American Conference on Bilingual Education* (pp. 63–72). Center for Applied Linguistics.

1976

- Bruck, M., Lambert, W. E., & Tucker, G. R. (1976). Alternative forms of immersion for second language teaching. *NABE*, 1(3), 33–48.
- Tucker, G. R., Hamayan, E., & Genesee, F. H. (1976). Affective, cognitive and social factors in second-language acquisition. *Canadian Modern Language Review*, 32(3), 214–226.

1977

- Bruck, M., Lambert, W. E., & Tucker, G. R. (1977). Cognitive consequences of bilingual schooling: The St. Lambert project through grade six. *Linguistics*, 15, 13–33.
- Hamayan, E., Genesee, F., & Tucker, G. R. (1977). Affective factors and language exposure in second language learning. *Language Learning*, 27(2), 225–241.
- Smith, P. M., Tucker, G. R., Taylor, D. M. (1977). Language, ethnic identity and intergroup relations: One immigrant group's reaction to language planning in Québec. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 283–306). Academic Press.
- Tucker, G. R. (1977). Some observations concerning bilingualism and second-language teaching in developing countries and in North America. In P. A. Hornby (Ed.), *Bilingualism: Psychological, social and educational implications* (pp. 141–146). Academic Press.
- Tucker, G. R. (1977). Can a second language be taught? In H. D. Brown, C. A. Yorio, & R. H. Crymes (Eds.), *Teaching and learning English as a Second Language: Trends in research and practice* (pp. 14–30). TESOL.

1978

- Acheson, P., d'Anglejan, A., de Bagheera, J., & Tucker, G. R. (1978). English as the second language in Quebec: A teacher profile. *McGill Journal of Education*, 13(2), 189–197.
- Genesee, F., Tucker, G. R., & Lambert, W. E. (1978). An experiment in trilingual education: Report 3. *Canadian Modern Language Review*, 34(3), 621–643.
- Genesee, F., Tucker, G. R., & Lambert, W. E. (1978). The development of ethnic identity and ethnic role taking skills in children from different school settings. *International Journal of Psychology*, 13(1), 39–57.
- Hamayan, E., Markman, B. R., Pelletier, S., & Tucker, G. R. (1978). Differences in performance in elicited imitation between French monolingual and English-speaking bilingual children. *IRAL: International Review of Applied Linguistics in Language Teaching*, 16(4), 330–339.
- Tucker, G. R., & Cziko, G. A. (1978). The role of evaluation in bilingual education. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1978* (pp. 423–446). Georgetown University Press.

1979

- Alatis, J. E., & Tucker, G. R. (Eds.). (1979). *Georgetown University round table on language and linguistics 1979*. Georgetown University Press.

- Cooper, R. L., Olshtain, E., Tucker, G. R., & Waterbury, M. (1979). The acquisition of complex English structures by adult native speakers of Arabic and Hebrew. *Language Learning*, 29(2), 255–275.
- Cowan, J. R., Light, R. L., Mathews, B. E., & Tucker, G. R. (1979). English teaching in China: A recent survey. *TESOL Quarterly*, 13(4), 465–482.
- Gray, T. C., & Tucker, G. R. (1979). Language policy and the delivery of social services in Canada and the United States. In J. E. Alatis & G. R. Tucker (Eds.), *Georgetown University round table on language and linguistics 1979* (pp. 61–75). Georgetown University Press.
- Tucker, G. R. (1979). Bilingual education: Some perplexing observations. *Educational Evaluation and Policy Analysis*, 1(5), 74–75.
- Tucker, G. R., & Sarofim, M. (1979). Investigating linguistic acceptability with Egyptian EFL students. *TESOL Quarterly*, 13(1), 29–39.

1980

- Hamayan, E., & Tucker, G. R. (1980). Language input in the bilingual classroom and its relationship to second language achievement. *TESOL Quarterly*, 14(4), 453–468.

1981

- Rhodes, N.C., Tucker, G. R., & Clark, J. L. D. (1981). *Elementary school foreign language instruction in the United States: Innovative approaches for the 1980's*. Center for Applied Linguistics.
- Tucker, G. R. (1981). Language issues for the 1980's: Examples from the United States. *Canadian Modern Language Review*, 37(3), 441–446.
- Tucker, G. R. (1981). Concluding thoughts. *Annual Review of Applied Linguistics*, 2, 254–258.
- Tucker, G. R. (1981). Bilingual education research in developing countries: Implications for the United States. *California Journal of Teacher Education*, 8(3), 1–14.
- Tucker, G. R. (1981). Social policy and second language teaching. In R. C. Gardner, & R. Kalin (Eds.), *A Canadian social psychology of ethnic relations* (pp. 77–92). Methuen.

1982

- Cziko, G. A., Lambert, W. E., Sidoti, N., & Tucker, G. R. (1982). Graduates of early immersion: Retrospective views of grade 11 students and their parents. In R. N. St. Clair, & H. Giles (Eds.), *The social and psychological contexts of language*. Lawrence Erlbaum Associates.
- Lambert, W. E., & Tucker, G. R. (1982). Graduates of early French immersion. In G. Caldwell & E. Waddell (Eds.), *The English of Quebec: From majority to minority status* (pp. 259–277). Institut Québécois de recherche sur la culture.

1983

- Tucker, G. R. (1983). Developing a research agenda for second language educators. In J. E. Alatis, H. H. Stern, & P. Strevens (Eds.), *Georgetown University round*

table on languages and linguistics 1983 (pp. 386–394). Georgetown University Press.

1984

Tucker, G. R. (1984). Toward the development of a language-competent American society. *International Journal of the Sociology of Language*, 45, 153–160.

Tucker, G. R. (1984). The future of language policy in education. In S. Shapson, & V. D'Oyley (Eds.), *Bilingual and multicultural education: Canadian perspectives* (pp. 143–153) Multilingual Matters.

1985

Tucker, G. R. (1985). Research directions and implications. *Annual Review of Applied Linguistics*, 6, 241–249.

Tucker, G. R., & Crandall, J. (1985). Innovative foreign language teaching in elementary schools. In P. H. Nelde (Ed.), *Methods in contact linguistic research*. Dummler.

1986

Tucker, G. R. (1986). Implications of Canadian research for promoting a language competent American society. In J. A. Fishman (Ed.), *The Fergusonian impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday Vol 2* (pp. 361–369). Mouton de Gruyter.

Tucker, G. R. (1986). Developing a language-competence American society. In D. Tannen, & J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1985* (pp. 263–274). Georgetown University Press.

1987

Tucker, G. R. (1987). Educational language policy in the Philippines: A case study. In P. H. Lowenberg (Ed.), *Georgetown University round table on languages and linguistics 1987* (pp. 331–341). Georgetown University Press.

1989

Tucker, G. R. (1989). Cognitive and social correlates of bilinguality. In R. Cooper & B. Spolsky (Eds.), *The influence of language on culture and thought: Essays in honor of Joshua A. Fishman's sixty-fifth Birthday* (pp. 101–111). Mouton de Gruyter.

Tucker, G. R., & Crandall, J. (1989). The integration of language and content instruction for language minority and language majority students. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1989* (pp. 39–50). Georgetown University Press.

1990

Crandall, J., & Tucker, G. R. (1990). Content-based instruction in second and foreign languages. In A. Padilla, H. Fairchild, & C. Valadez (Eds.), *Foreign language education: Issues and strategies* (pp. 187–200). Sage.

- Tucker, G. R. (1990). Second language education: Issues and perspectives. In A. Padilla, H. Fairchild, & C. Valadez (Eds.), *Foreign language education: Issues and strategies* (pp. 13–21). Sage.
- Tucker, G. R. (1990). An overview of applied linguistics. In M. A. K. Halliday, J. Gibbons, & H. Nicholas (Eds.), *Learning, keeping, and using language Vol I* (pp. 1–6). John Benjamins.
- Tucker, G. R. (1990). Cognitive and social correlates and consequences of additive bilinguality. In J. E. Alatis (Eds.), *Georgetown University round table on languages and linguistics 1990* (pp. 90–101). Georgetown University Press.

1991

- Padilla, A. M., Lindholm, K. J., Chen, A., Durán, R., Hakuta, K., Lambert, W., & Tucker, G. R. (1991). The English-only movement: Myths, reality, and implications for psychology. *American Psychologist*, *46*(2), 120–130.
- Tucker, G. R. (1991). Developing a language-competent American society: The role of language planning. In A. G. Reynolds (Ed.), *Bilingualism, multiculturalism, and second language learning: The McGill Conference in Honour of Wallace E. Lambert* (pp. 65–79). Psychology Press.

1993

- Tucker, G. R. (Ed.). (1993). *Policy and practice in the education of culturally and linguistically diverse students: Views from language educators*. Alexandria, VA: TESOL.
- Tucker, G. R. (1993). Language learning for the 21st century: challenges of the North American Free Trade Agreement. *Canadian Modern Language Review*, *50*(1), 165–172.

1994

- Donato, R., Antonek, J. L., & Tucker, G. R. (1994). A multiple perspectives analysis of a Japanese FLES program. *Foreign Language Annals*, *27*(3), 365–378.
- Tucker, G. R. (1994). Concluding thoughts: Language planning issues for the coming decade. *Annual Review of Applied Linguistics*, *14*, 277–283.

1995

- Tucker, G. R., & Donato, R. (1995). Developing a second language research component within a teacher education program. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1995* (pp. 453–470). Georgetown University Press.

1996

- Donato, R., Antonek, J. L., & Tucker, G. R. (1996). Monitoring and assessing a Japanese FLES program: Ambiance and achievement. *Language Learning*, *46*(3), 497–528.
- Tucker, G. R. (1996). Some thoughts concerning innovative language education programmes. *Journal of Multilingual and Multicultural Development*, *17*(2–4), 315–320.

Tucker, G. R., Donato, R., & Antonek, J. L. (1996). Documenting growth in a Japanese FLES program. *Foreign Language Annals*, 29(4), 539–550.

1997

Paulston, C. B., & Tucker, G. R. (Eds.). (1997). *The early days of sociolinguistics: Memories and reflections*. Summer Institute of Linguistics.

Tucker, G. R. (1997). Developing a language-competent American society: Implications of the English-only Movement. In T. Bongaerts & K. de Bot (Eds.), *Perspectives on foreign language policy: Studies in honor of Theo van Els* (pp. 89–98). John Benjamins.

Tucker, G. R. (1997). Multilingualism and language contact: An introduction. *Annual Review of Applied Linguistics*, 17, 3–10.

Tucker, G. R. (1997). The development of sociolinguistics as a field of study: Concluding observations. In C. B. Paulston & G. R. Tucker (Eds.), *The early days of sociolinguistics: Memories and reflections* (pp. 317–324). Summer Institute of Linguistics.

Tucker, G. R., & Corson, D. (Eds.). (1997). *Encyclopedia of language and education Vol 4: Second language education*. Kluwer Academic Publishers.

1998

Tucker, G. R. (1998). Priorities for a research agenda for early language learning. In M. Met (Ed.), *Critical issues in early second language learning: building for our children's future* (pp. 308–339). Addison-Wesley.

Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 3–15). Multilingual Matters.

Tucker, G. R., & Donato, R. (1998). Designing and implementing an innovative foreign language program: Reflections from a school district-university partnership. *Studia Anglica Posnaniensia: International review of English Studies*, 33, 451–461.

2000

Antonek, J., Donato, R., Tucker, G. R. (2000). Differential linguistic development of Japanese language learners in elementary school. *Canadian Modern Language Review*, 57(2), 325–351.

Donato, R., Tucker, G. R., Wudthayagorn, J., & Igarashi, K. (2000). Converging evidence: Attitudes, achievements, and instruction in the later years of FLES. *Foreign Language Annals*, 33(4), 377–393.

Tucker, G. R. (2000a). The applied linguist, school reform, and technology: challenges and opportunities for the coming decade. *CALICO Journal*, 17(2), 197–221.

Tucker, G. R. (2000b). Concluding thoughts: Applied linguistics at the juncture of millennia. *Annual Review of Applied Linguistics*, 20, 241–249.

Tucker, G. R. (2000c). Precision, elegance and simplicity: Perspectives on TESOL and art. *Journal of the Imagination in Language Learning*, 5, 24–26.

Tucker, G. R. (2000d). The importance of longitudinal research for language educators. In M. L. S., Bautista, T. A., Llamzon, & B. P., Sibayan (Eds.), *Parangal nang Brother Andrew: Festschrift for Andrew Gonzalez on his sixtieth birthday* (pp. 161–169). Linguistic Society of the Philippines.

2001

Tucker, G. R. (2001a). A global perspective on bilingualism and bilingual education. In J. E. Alatis & A.-H. Tan (Eds.), *Georgetown University round table on languages and linguistics 1999* (pp. 332–340). Georgetown University Press.

Tucker, G. R. (2001b). Age of beginning instruction. *TESOL Quarterly*, 35(4), 597–598.

Tucker, G. R., Donato, R., & Murday, K. (2001). The genesis of a district-wide Spanish FLES program. In R. L. Cooper, E. Shohamy, & J. Walters (Eds.), *New perspectives and issues in educational language policy: In honour of Bernard Dov Spolsky* (pp. 235–259). John Benjamins.

2002

Chinen, K., & Tucker, G. R. (2002). Saturday-school participation, ethnic identity and Japanese language development. *Heritage Language Journal*, 1(1).

Igarashi, K., Wudthayagorn, J., Donato, R., & Tucker, G. R. (2002). What does a novice look like? Describing the grammar and discourse of young learners of Japanese. *Canadian Modern Language Review*, 58(4), 526–554.

2003

Chinen, K., Donato, R., Igarashi, K., & Tucker, G. R. (2003). Looking across time: Documenting middle school Japanese FLES students' attitudes, literacy and oral proficiency. *Learning Languages*, 8(2), 4–10.

Paulston, C. B., & Tucker, G. R. (Eds.). (2003). *Sociolinguistics: The essential readings*. Blackwell.

Tucker, G. R., & Donato, R. (2003). Implementing a district-wide foreign language program: A case study of acquisition planning and curricular innovation. In D. Tannen, & J. E. Alatis (Eds.), *Georgetown University round table on languages and linguistics 2001* (pp. 178–195). Georgetown University Press.

Tucker, G. R. (2003). Language contact and change: Summary observations. *Annual Review of Applied Linguistics*, 23, 243–249.

2004

Tucker, G. R. (2004). Learning other languages: The case for promoting bilinguality within an international community. In D. Resnick, & D. Scott (Eds.), *The innovative university*. Carnegie Mellon University Press.

2005

Chinen, K., & Tucker, G. R. (2005). Heritage language development: Understanding the roles of ethnic identity and Saturday school participation. *Heritage Language Journal*, 3(1), 27–59.

Dominguez, R., Tucker, G. R., & Donato, R. (2005). Documenting curricular reform: Innovative foreign language education in elementary school. In

P. Bruthiaux, D. Atkinson, W. G. Eggington, W. Grabe, & V. Ramanathan (Eds.), *Directions in applied linguistics: Essays in honor of Robert B. Kaplan* (pp. 56–71). Multilingual Matters.

Mitsui, A., Morimoto, Y., Tucker, G. R., & Donato, R. (2005). “Intelligent” errors: Kanji writing as meaning making for Japanese FLES learners. *Learning Languages*, 11(1), 5–14.

Tucker, G. R. (2005). Innovative language education programmes for heritage language students: The special case of Puerto Ricans? *International Journal of Bilingual Education and Bilingualism*, 8(2–3), 188–195.

2006–2010

Chenoweth, N. A., Jones, C. M., & Tucker, G. R. (2006). Language Online: Principles of design and methods of assessment. In R. P. Donaldson & M. A. Haggstrom (Eds.), *Changing language education through CALL* (pp. 149–168). Routledge.

Chinen, K., & Tucker, G. R. (2006). Heritage language development: Understanding the roles of ethnic identity, schooling and community. In K. Kondo-Brown (Ed.), *Heritage language development: Focus on East Asian immigrants* (pp. 89–126). John Benjamins.

Donato, R., & Tucker, G. R. (2007). K-12 language learning and foreign language educational policy: A school-based perspective. *Modern Language Journal*, 91(2), 256–258.

Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.

Mitsui, A., Donato, R., & Tucker, G. R. (2007). What can you remember? Recall of Japanese after a one-year hiatus. *Learning Languages*, 12(2), 8–12.

Mitsui, A., Donato, R., & Tucker, G. R. (2009). Introducing writing activities into the Japanese-as-a-foreign-language (JFL) classroom: A teacher-researcher collaboration. *NECTFL Review*, 64, 5–17.

Pessoa, S. Hendry, H., Donato, R., Tucker, G. R., & Lee, H. (2007). Content-based instruction in the foreign language classroom: A discourse perspective. *Foreign Language Annals*, 40(1), 102–121.

Tucker, G. R. (2008). Learning other languages: The case for promoting bilingualism in our educational system. In D. M. Brinton, O. Kagan, & S. Bauckus (Eds.), *Heritage Language education: A new field emerging* (pp. 39–52). Routledge.

Tucker, G. R. & Dubiner, D. (2009). Concluding thoughts: Does the immersion pathway lead to multilingualism? In T. W. Fortune & D. J. Tedick (Eds.), *Pathways to multilingualism: Evolving perspectives on immersion education* (pp. 267–277). Multilingual Matters.

2011–2015

Donato, R., Tucker, G. R., & Hendry, H. (2014). Developing professional identities in applied linguistics: From doctoral study to professional practice. In Y. L. Cheung, S. B. Said, & K. Park (Eds.), *Advances and current trends in language teacher identity research* (pp. 217–234). Routledge.

- Li, S., & Tucker, G. R. (2013). A survey of the US Confucius Institutes: Opportunities and challenges in promoting Chinese language and culture education. *Journal of the Chinese Language Teachers Association*, 48(1), 29–54.
- Nair, I., Norman, M., Tucker, G. R., & Burkert, A. (2012). The challenge of global literacy: An ideal opportunity for liberal professional education. *Liberal Education*, 98(1), 56–61.

2016–2018

- Ke, S., & Tucker, G. R. (2016). Unleavened cakes: L2 Chinese Education for South Asian Students in Hong Kong. *K-12 Chinese Language Teaching*. <https://clta-us.org/wp-content/uploads/2015/09/Unleavened-Cakes-L2-Chinese-Education-for-South-Asian-Students-in-Hong-Kong.pdf>
- Polansky, S., & Tucker, G. R. (2018). Modern languages majors in the twenty-first century: Broadening disciplinary frames of reference and global awareness. *ADFL Bulletin*, 45(1), 139–149.
- Tucker, G. R. (2018). Summary and concluding observations. In J. Crandall & K. M. Bailey (Eds.), *Global perspectives on language education policies* (pp. 219–225). Routledge and TIRF.

References

- Ackerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research*, 81(2), 132–169.
- Atkinson, D. (2002). Toward a sociocognitive approach to second language acquisition. *Modern Language Journal*, 86(4), 525–545.
- Barnett, J. E. (2008). Mentoring, boundaries, and multiple relationships: Opportunities and challenges. *Mentoring & Tutoring: Partnership in Learning*, 16(1), 3–16.
- Bruck, M., Lambert, W. E., & Tucker, G. R. (1974). Bilingual schooling through the elementary grades: The St. Lambert project at grade 7. *Language Learning*, 24(2), 183–204.
- Chinen, K., Donato, R., Igarashi, K., & Tucker, G. R. (2003). Looking across time: Documenting middle school Japanese FLES students' attitudes, literacy and oral proficiency. *Learning Languages*, 8(2), 4–10.
- Cowan, J. R., Light, R. L., Mathews, B. E., & Tucker, G. R. (1979). English teaching in China: A recent survey. *TESOL Quarterly*, 13(4), 465–482.
- Crandall, J., & Tucker, G. R. (1990). Content-based instruction in second and foreign languages. In A. Padilla, H. Fairchild, & C. Valadez (Eds.), *Foreign language education: Issues and strategies* (pp. 187–200). Sage.
- Donato, R. (2013). In G. R. Tucker & I. C. A. Chapelle (Eds.), *The Encyclopedia of applied linguistics*. Wiley-Blackwell.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.
- Donato, R., Tucker, G. R., & Hendry, H. (2014). Developing professional identities in applied linguistics: From doctoral study to professional practice. In Y. L. Cheung, S. B. Said, & K. Park (Eds.), *Advances and current trends in language teacher identity research* (pp. 217–234). Routledge.
- Engeström, Y., Engeström, R., & Kärkkäinen, M. (1995). Polycontextuality and boundary crossing in expert cognition: Learning and problem solving in complex work activities. *Learning and Instruction*, 5, 319–336.

- García, O., & Li, W. (2014). *Translanguaging: Language, bilingualism, and education*. Palgrave Macmillan.
- Hall, J. K. (1997). A consideration of SLA as a theory of practice: A response to Firth and Wagner. *Modern Language Journal*, 97, 301–306.
- Hulstijn, J. H., Young, R. F., Ortega, L., & Bigelow, M. (2014). Bridging the gap: Cognitive and social approaches to research in second language learning and teaching. *Studies in Second Language Acquisition*, 36(3), 361–421.
- Igarashi, K., Wudthayagorn, J., Donato, R., & Tucker, G. R. (2002). What does a novice look like? Describing the grammar and discourse of young learners of Japanese. *Canadian Modern Language Review*, 58(4), 526–554.
- Jacobson, M. J., Levin, J. A., & Kapur, M. (2019). Education as a complex system: Conceptual and methodological implications. *Educational Researcher*, 48(2), 112–119.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.
- King, K. A., & Mackey, A. (2016). Research methodology in second language studies: Trends, concerns, and new directions. *Modern Language Journal*, 100(S1), 209–227.
- Kostoulas, A. (Ed.). (2019). *Challenging boundaries in language education*. Springer.
- Kramsch, C. (2008). Ecological perspectives on foreign language education. *Language Teaching*, 41(3), 389–408.
- Kumaravadivelu, B. (2003). *Beyond methods: Macrostrategies for language teaching*. Yale University Press.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Newbury House Publishers.
- Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics? *Language Teaching*, 45(2), 202–214.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Mitsui, A., Donato, R., & Tucker, G. R. (2007). What can you remember? Recall of Japanese after a one-year hiatus. *Learning Languages*, 12(2), 8–12.
- Nair, I., Norman, M., Tucker, G. R., & Burkert, A. (2012). The challenge of global literacy: An ideal opportunity for liberal professional education. *Liberal Education*, 98(1), 56–61.
- Ortega, L. (2010). The bilingual turn in SLA. Plenary delivered at the Annual Conference of the American Association for Applied Linguistics. Atlanta, GA, March 6–9.
- Pennycook, A. (2001). *Critical applied linguistics: A critical introduction*. Lawrence Erlbaum Associates.
- Polansky, S., & Tucker, G. R. (2018). Modern languages majors in the 21st century: Broadening disciplinary frames of reference and global awareness. *ADFL Bulletin*, 45(1), 139–149.
- Riazi, A. M., & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues and challenges. *Language Teaching*, 47(2), 135–173.
- Suchman, L. (1994). Working relations of technology production and use. *Computer Supported Cooperative Work*, 2, 21–39.
- The Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *Modern Language Journal*, 100(S1), 19–47.
- Tucker, G. R. (1991). Developing a language-competent American society: The role of language planning. In A. G. Reynolds (Ed.), *Bilingualism, multiculturalism, and second language learning: The McGill conference in honour of Wallace E. Lambert* (pp. 65–79). Psychology Press.
- Tucker, G. R. (1997). Developing a language-competent American society implications of the English-only movement. In T. Bongaerts & K. de Bot (Eds.), *Perspectives on foreign language policy: Studies in honor of Theo van Els* (pp. 89–98). John Benjamins.
- Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 3–15). Multilingual Matters.

- Tucker, G. R. (2000a). The applied linguist, school reform, and technology: Challenges and opportunities for the coming decade. *CALICO Journal*, 17(2), 197–221.
- Tucker, G. R. (2000b). Concluding thoughts: Applied linguistics at the juncture of millennia. *Annual Review of Applied Linguistics*, 20, 241–249.
- Tucker, G. R. (2000c). Precision, elegance and simplicity perspectives on TESOL and art. *Journal of the Imagination in Language Learning*, 5, 24–26.
- Tucker, G. R. (2001). A global perspective on bilingualism and bilingual education. In J. E. Alatis & A.-H. Tan (Eds.), *Georgetown University round table on languages and linguistics 1999* (pp. 332–340). Georgetown University Press.
- Tucker, G. R., & Corson, P. (Eds.). (1997). *Encyclopedia of language and education: Second language education*. Springer.
- Tucker, G. R., & Crandall, J. (1989). The integration of language and content instruction for language minority and language majority students. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1989* (pp. 39–50). Georgetown University Press.
- Tucker, G. R., Donato, R., & Antonek, J. L. (1996). Documenting growth in a Japanese FLES program. *Foreign Language Annals*, 29(4), 539–550.
- Van Lier, L. (2004). *The ecology and semiotics of language learning: A sociocultural perspective*. Kluwer.
- Zuengler, J., & Miller, E. R. (2006). Cognitive and sociocultural approaches: Two parallel SLA worlds? *TESOL Quarterly*, 40(1), 35–58.

Part I

Language Learning and Development

Introduction and the Tuckerian Impact

This part focuses on the learner and on language learning and development. To understand the broader issues that define the focus of the other parts, it is essential that we cross boundaries to probe into the learner and learning and develop a solid understanding about such fundamental issues as mechanisms and processes of learning; social and psychological factors in learning; cognitive and social “consequences” of language learning and bilinguality; and, more broadly, language learning, schooling, and educational achievement.

This thematic area forms an important part of Dick’s scholarship and contribution to applied linguistics. With collaborators and graduate students, Dick published numerous articles that examined processes of language learning and development and the personal, social, and educational “consequences” of language learning and bilinguality. Sometimes, the studies reported were contextualized in a broader agenda such as program implementation and evaluation (e.g., Antonek et al., 2000; Chinen et al., 2003; Genesee et al., 1978; Lambert et al., 1973; see also Part III) or comparison of instructional methods and classroom processes (e.g., Hamayan & Tucker, 1980; Tucker et al., 1969; see also Part II). Other times, they were purposefully designed to investigate a very specific issue in language learning (e.g., Chinen & Tucker, 2005; Scott & Tucker, 1974). These publications, collectively and individually, exemplify the crossing of boundaries between languages, learners, contexts, theoretical perspectives, and/or methodological approaches (see Fig. 1, Chap. 1).

Specifically, these publications crossed boundaries of target languages and language backgrounds of learners, including but not limited to, Arabic- and Hebrew-speaking (e.g., Cooper et al., 1979; Scott & Tucker, 1974; Tucker & Sarofim, 1979), French-speaking (e.g., d’Anglejan & Tucker, 1975), and Filipino-speaking learners of English (Tucker et al., 1970); English-speaking learners of French (Bruck et al., 1974; Tucker et al., 1969), Spanish (Donato & Tucker, 2010), and Japanese (Antonek et al., 2000; Igarashi et al., 2002; Tucker et al., 1996); as well as heritage learners of

Japanese (e.g., Chinen & Tucker, 2002, 2005). Boundary crossing was sometimes also manifested in the design of individual studies that involved, for example, comparing L1 and L2 users of a language, such as French (Bruck et al., 1974; Lambert et al., 1973) and Japanese (Mitsui et al., 2005), and exploring L1 influence on L2 learning and development (Scott & Tucker, 1974).

Other than the linguistic diversity, Dick's publications also involved other learner diversities, such as young (e.g., Bruck et al., 1974) or adult learners (e.g., d'Anglejan & Tucker, 1975; Saegert et al., 1974; Scott & Tucker, 1974; Tucker et al., 1969); and learners in diverse social and educational settings, including Canada (d'Anglejan & Tucker, 1975; Tucker et al., 1969, 1976;), Egypt (Cooper et al., 1979; Tucker & Sarofim, 1979), Israel (Cooper et al., 1979), Lebanon (Scott & Tucker, 1974), Japan (Mitsui et al., 2005), the Philippines (Gatbonton & Tucker, 1971; Tucker et al., 1970), and the United States (Chinen & Tucker, 2002, 2005; Tucker et al., 1996).

Dick is a strong advocate for crossing disciplinary, paradigmatic, and methodological boundaries (and for educating graduate students and junior researchers to cross these boundaries as well). His work underscores the importance of longitudinal research and developmental perspectives on language learning (and language education in general) (see Tucker, 2000a, 2000b, 2000c). As he argued,

we should not encourage the importation of any one relatively restricted research paradigm or tradition" and "we should encourage our students to seek the broadest possible training in qualitative and ethnographic as well as in quantitative techniques and [...] we should then work to ensure that we all use the tools that are most appropriate for the questions we are asking from among the broad array of techniques and procedures currently available to us (Tucker, 2000b, p. 207).

Crossing paradigmatic and methodological boundaries is also clearly exemplified in Dick's research with collaborators and graduate students. For example, Dick's studies on language learning and development were influenced by linguistic and psycholinguistic (e.g., Bruck et al., 1974; Cohen et al., 1967; Davine et al., 1971; Tucker et al., 1969) as well as socio-psychological and sociolinguistic perspectives (e.g., Genesee et al., 1978; Tucker et al., 1976). They also followed diverse approaches, designs, and methods, ranging from quantitative, experimental and factorial designs, and correlation-based research (e.g., Gatbonton & Tucker, 1971; Genesee et al., 1975; Lambert et al., 1973; Saegert et al., 1974) to qualitative interviews, linguistic analysis, and case comparisons (e.g., Igarashi et al., 2002; Mitsui et al., 2005). Sometimes, methods and sources of data were purposefully mixed to understand and interpret language development and the consequences of language learning and bilinguality (e.g., Antonek et al., 2002; Mitsui et al., 2005; Scott & Tucker, 1974; Tucker et al., 1996).

References

- Antonek, J., Donato, R., & Tucker, G. R. (2000). Differential linguistic development of Japanese language learners in elementary school. *Canadian Modern Language Review*, 57(2), 325–351.
- Bruck, M., Lambert, W. E., & Tucker, G. R. (1974). Bilingual schooling through the elementary grades: The St. Lambert Project at grade 7. *Language Learning*, 24(2), 183–204.
- Chinen, K., & Tucker, G. R. (2002). Saturday-school participation, ethnic identity and Japanese language development. *Heritage Language Journal*, 1(1). Retrieved from <https://www.international.ucla.edu/ccs/article/3902>
- Chinen, K., & Tucker, G. R. (2005). Heritage language development: Understanding the roles of ethnic identity and Saturday school participation. *Heritage Language Journal*, 3(1), 27–59.
- Chinen, K., Donato, R., Igarashi, K., & Tucker, G. R. (2003). Looking across time: Documenting middle school Japanese FLES students' attitudes, literacy and oral proficiency. *Learning Languages*, 8(2), 4–10.
- Cohen, S. P., Tucker, G. R., & Lambert, W. E. (1967). The comparative skills of monolinguals and bilinguals in perceiving phoneme sequences. *Language and Speech*, 10(3), 159–168.
- Cooper, R. L., Olshtain, E., Tucker, R. G., & Waterbury, M. (1979). The acquisition of complex English structures by adult native speakers of Arabic and Hebrew. *Language Learning*, 29(2), 255–275.
- d'Anglejan, A., & Tucker, G. R. (1975). The acquisition of complex English structures by adult learners. *Language Learning*, 25(2), 281–296.
- Davine, M., Tucker, G. R., & Lambert, W. E. (1971). The perception of phoneme sequences by monolingual and bilingual elementary school children. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, 3(1), 72–76.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.
- Gatbonton, E. C., & Tucker, G. R. (1971). Cultural orientation and the study of foreign literature. *TESOL Quarterly*, 5(2), 137–143.
- Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communication skills of bilingual children. *Child Development*, 46(4), 1010–1014.
- Genesee, F., Tucker, G. R., & Lambert, W. E. (1978). The development of ethnic identity and ethnic role taking skills in children from different school settings. *International Journal of Psychology*, 13(1), 39–57.
- Hamayan, E., & Tucker, G. R. (1980). Language input in the bilingual classroom and its relationship to second language achievement. *TESOL Quarterly*, 14(4), 453–468.
- Igarashi, K., Wudthayagorn, J., Donato, R., & Tucker, G. R. (2002). What does a novice look like? Describing the grammar and discourse of young learners of Japanese. *Canadian Modern Language Review*, 58(4), 526–554.
- Lambert, W. E., Tucker, G. R., & d'Anglejan, A. (1973). Cognitive and attitudinal consequences of bilingual schooling. *Journal of Educational Psychology*, 65(2), 141–159.
- Mitsui, A., Morimoto, Y., Tucker, G. R., & Donato, R. (2005). "Intelligent" errors: Kanji writing as meaning making for Japanese FLES learners. *Learning Languages*, 11(1), 5–14.
- Saegert, J., Scott, S., Perkins, J., & Tucker, G. R. (1974). A note on the relationship between English proficiency, years of language study and medium of instruction. *Language Learning*, 24(1), 99–104.
- Scott, M. S., & Tucker, G. R. (1974). Error analysis and English-language strategies of Arab students. *Language Learning*, 24(1), 69–97.
- Tucker, G. R. (2000a). The applied linguist, school reform, and technology: Challenges and opportunities for the coming decade. *CALICO Journal*, 17(2), 197–221.
- Tucker, G. R. (2000b). Concluding thoughts: Applied linguistics at the juncture of millennia. *Annual Review of Applied Linguistics*, 20, 241–249.

- Tucker, G. R. (2000c). The importance of longitudinal research for language educators. In M. L. S. Bautista, T. A. Llamzon, & B. P. Sibayan (Eds.), *Parangal nang Brother Andrew: Festschrift for Andrew Gonzalez on his sixtieth birthday* (pp. 161–169). Linguistic Society of the Philippines.
- Tucker, G. R., & Sarofim, M. (1979). Investigating linguistic acceptability with Egyptian EFL students. *TESOL Quarterly*, 13(1), 29–39.
- Tucker, G. R., Donato, R., & Antonek, J. L. (1996). Documenting growth in a Japanese FLES program. *Foreign Language Annals*, 29(4), 539–550.
- Tucker, G. R., Hamayan, E., & Genesee, F. H. (1976). Affective, cognitive and social factors in second-language acquisition. *Canadian Modern Language Review*, 32(3), 214–226.
- Tucker, G. R., Lambert, W. E., & Rigault, A. (1969). Students' acquisition of French gender distinctions: A pilot investigation. *IRAL - International Review of Applied Linguistics in Language Teaching*, 7(1), 51–55.
- Tucker, G. R., Otones, F. T., & Sibayan, B. P. (1970). An alternate days approach to bilingual education. In J. A. Atlatis (Ed.), *Georgetown University round table on languages and linguistics 1970* (pp. 281–295). Georgetown University Press.

“I Want the Next Experience”: Israeli Adult Native Bilinguals Tell the Story of Their Childhood Bilinguality



Deborah Dubiner

Abstract This chapter focuses on a retrospective report of the language experiences of native bilingual Israelis who were born to immigrant parents around independence (1948). It does so by examining narrative life stories of adult Israelis who reconstruct, and reflect on, the impact of language(s) in their lives. Growing up in a period of transition from the age-known identity of a stateless Diaspora Jew to the brand-new Israeli Jewish identity, children had the tacit responsibility of helping revive the Hebrew language. The interviewees describe their exposure to a language other than Hebrew at home and the emotional and psychological load of mediating between their parents' and their own identities. These bilingual adults report on the journey of boundary crossing entailed by this reality. The narratives illuminate the speakers' perceptions and their bilinguality during a period when speaking Hebrew was a cornerstone in shaping the Israeli ethos and imperative in identity formation of the new State and of its citizens. By underscoring each child's mediation of their own bilinguality, and their subsequent engagement with the language of wider communication, this chapter illustrates how boundary crossing in research can inform language education and language education research.

Keywords Language ideologies · Childhood bilingualism · Bilingualism and identity · Language choice · Linguistic mudes

1 Introduction

In twenty-first-century multicultural environments worldwide, it is not uncommon to hear of families whose home language is different from the dominant language of the community in which they live. Clearly, Spanish is spoken in San Diego, Russian

D. Dubiner (✉)
Oranim College of Education, Kiryat Tiv'on, Israel
e-mail: dvora_d@oranim.ac.il

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_2

in Tel Aviv, Arabic in Paris, Turkish in Berlin, and the list goes on. Sometimes, these languages may cross the boundaries of the home and be heard outside; other times, family language policy, often dictated by children, may determine that the home language will stay as such – in the home. Language maintenance is often chosen as a way of keeping an ethnic or cultural identity and link with the culture of origin. In these cases, language maintenance may occur even if individuals function actively in the dominant language in the wider community (Myers-Scotton, 2005).

The purpose of the present chapter is to explore language use and identity-building of in the personal, familial, and national realms in post-independence Israel. An examination of narratives of adult native bilingual Israelis who were born to immigrant parents in the first years after the establishment of the State of Israel (1948) might shed light on the nature of boundary crossing by children in a period when speaking the societal language was vital for identity formation. Children then were tacitly given the responsibility not only for their own identity but for nothing less than an entire country's national language revival.

2 Theoretical Background

2.1 *Language and National Identity: The Case of Hebrew*

Multilingualism is pervasive in many regions around the globe and believed to be the unmarked case worldwide (Tucker, 1998). When functioning in different languages, individuals constantly negotiate the identities that come about through the languages they speak. In this respect, research has shown a conclusive link between language use and identity (see, for example, Dubiner, 2018; First & Avraham, 2010; Gu, 2011; Leeman et al., 2011; Mu, 2015; Pavlenko & Blackledge, 2004; Wodak, 2012; Zentella, 1997). Hamers and Blanc (2000) further indicate that not only do language and ethnolinguistic identity develop simultaneously, but there exists a reciprocal relationship between the two.

Undeniably, language is one of the characteristics that defines who we are, and this identity may be assigned to us by ourselves or by the society at large. It is a symbolic component of an individual's identity; it is also one of the “emblematic features” (Creese & Blackledge, 2015, p. 23) that determines group membership. Language choice and language ideology are interrelated (Pavlenko & Blackledge, 2004) and the choice bears much symbolic significance. The sociolinguistic context, with its ideologies and attendant symbols, is therefore a key defining factor in one's identity.

In this sense, the connection between language and national identity is multifaceted and investigated by sociolinguists from different angles. One of the perspectives offered in the literature is that language is “a major element in the development of an ethnic community's political consciousness and a tool of state-building ...” (Safran, 1992, p. 397). Varied contexts give birth to diverse combinations of minority language use in the face of the official language, the language of wider

communication, or the dominant language of a country, which is not always the dominant language of the population. In the case of Israel, after independence, Hebrew was seen as a symbol, and a builder, of a national unity. The use of other languages was perceived as a threat not only to the language, but to the nation itself. Safran (1992) claims that Hebrew was deliberately used a tool of state-building and a way of spreading the above-mentioned ideology – that of Hebrew as a national language of the new State. The author goes on to summarize that “it is difficult to dissociate the development of the modern Jewish national consciousness (Zionism) and the subsequent building of the Israeli polity from the Hebrew language” (p. 405).

2.2 Immigration to Israel, Language Maintenance, and the Revival of Hebrew

An obvious by-product of immigration is language maintenance, a topic largely discussed in sociolinguistic research. It refers to “keeping a language vital within a given community or region” (Hall et al., 2017, p. 113). Conversely, language shift (a shift from the home language to the dominant/target language) might occur when younger generations become monolinguals of the dominant language of the society in which they live (Fishman, 2013). Integration ideologies might motivate language practices that result in language shift; yet, as Fishman claims, this language shift can be reversed when families and their immediate community engage in child-adult personal connection processes that promote language use in a natural way. These “bonds of intergenerational and spontaneous affect, intimacy, identity, and loyalty” are pivotal in the transmission of home languages down to the next generations (Fishman, 2013, p. 493). Nevertheless, language maintenance and language shift occur and are dictated by individuals and by the way they *interpret* social change taking place in the community (Li, 2000). Language maintenance or language shift will depend on a plethora of factors, including “sociolinguistic or emotional pressures” (Sevinç, 2016), and, in fact, the way speakers interpret them or their importance.

In the late nineteenth century, after generations of being a stateless people and leading multilingual lives in the Diaspora, Jews began to “return to Zion” (Spolsky & Shohamy, 1999, p. 13). They were motivated by a fierce ideology and intense will of having a country, a place the Jewish people could call theirs. Before the proclamation of independence in 1948, waves of Jewish immigrants settled in the country and added to the extant Jewish population and the then-majority of local Arab residents. More specifically, between 1882 and 1947 about half a million Jews joined the 24,000 who already lived there (Neuman, 1999). From 1919 to the end of the British mandate in 1948, the Jewish population of Israel jumped from about 55,000–56,000 to 650,000 people (Lissak, 2009). After independence from Britain in 1948, waves of mass immigration occurred, especially in the first years of statehood. Yet Jewish immigrants continued to constantly move to Israel from all over

the world after the large waves such as those of the 1950s and the 1990s. As a group prone to assimilation (Smootha, 2008) and due to the local “melting pot” policy, Jewish immigrants often did not maintain the language of their country of origin.

From the middle of the nineteenth century, fierce and adamant efforts were made to stipulate Hebrew as the spoken language of Israel. Its revival, to which Fishman (2000) referred as nothing short of “miraculous” (p. 215), involved an encompassing abandonment of all other home languages. As a result of this ideology, Hebrew indeed ceased to be used solely in liturgical and religious contexts and became a spoken language among Jews in Israel and, after independence, the official language of the country, alongside Arabic.¹

Growing up in a period of transition from the age-old identity of a stateless Diaspora Jew to the brand-new Israeli Jewish identity, children, with their loyalty to this *modern Israeli*, were a key factor in helping revive the Hebrew language. As a national effort to revive the language, Jewish immigrants were expected to speak solely Hebrew and abandon their home languages. The narratives in the present study describe the participants’ exposure to a language other than Hebrew at home during childhood. They also depict the emotional load of mediating between their parents’ and their own identities.

2.3 *Bilingual Identity and Boundary Crossing*

The use of different languages by multilinguals bears with it much sociolinguistic significance. Language choice and language attitude are both a cause and a result of the constant encounter of each bi/multilingual individual with the realities present in the context of each of his or her languages. Language choice signifies group membership or a denial of group membership. In fact, Giles and Byrne (1982) point to a correlation between the strength of in-group boundaries and the level of assimilation into the out-group – or the crossing of those boundaries. Pavlenko and Blackledge (2004) stress that each particular situation “sets out the boundaries for particular *identity options*” (p. 25, emphasis mine), suggesting that a bilingual’s negotiation of identities implies a negotiation of boundaries as well. In her study of Verlan, a variety of French spoken by marginalized youth in suburban Paris, Doran (2004) points to language choice (standard French vs. Verlan) as a tool for expression of hybrid identities and marking of the boundaries between an ethnic community and the perceived bourgeoisie of the dominant group. These minority youths switched between languages (the heritage language, Verlan, at home and standard French) as a way of navigating their identities. They built boundaries by means of

¹ In 2018, Arabic lost its status as an official language and received the title of “a language with a special status” in Israel. As of now, this change has no practical effects on the citizens. For example, Arabic-speaking children continue to receive education in their mother tongue, and government documents continue to be accessible in Arabic.

their sociolect to determine a certain group identity, only to cross them later, enacting chosen identities as dictated by convenience or necessity.

Investigating bilinguals' boundary crossing can benefit language educators and language education researchers in their attempt to increase awareness of language learners' experiences in a variety of contexts. Taking notice of boundary crossing that children and adolescents experience may facilitate the guidance and support of bilinguals and incipient bilinguals in their navigation of complementing or conflicting identities. The present study engages in this inquiry by examining narratives of bilingual childhood experiences as described in the methodology below.

3 Methodology

When people tell a listener their life stories, they reconstruct and make sense of their memories regarding the events they are relating. Through narratives, speakers also explore and perhaps even mold their own identities (De Fina, 2015). Regarding immigration-related narratives, by telling their stories, narrators are able to scrutinize past experiences. This may aid bilingual individuals in better grasping life changes, for they facilitate the exploration of “both the pain and the value inherent in building an identity in a new language and culture” (Menard-Warwick, 2004, p. 297). The role of the researcher is to locate and identify elements in narratives that point to special or meaningful events that, together with special or meaningful events in other speakers' lives, are braided into a whole that might enable understanding of critical events in a community.

The most salient advantage of narrative analysis might be that it constitutes a unique tool to enter a narrator's internal world, “inaccessible to experimental methodologies”, thus providing “the insider's view of the processes of language learning, attrition, and use” (Pavlenko, 2007, pp. 164–165). Through the lens of first-hand actors, we are presented with bricks that build the sociocultural context we are trying to understand. Additionally, only through personal narratives are we able to learn about the personal perspectives regarding experiences and events that escape the broader scope of “historic and diachronic sociolinguistic research” (Pavlenko, 2007, p. 165).

The study presented in this chapter is based on the analyses of narratives of five adult Israeli bilinguals who were born between the years 1951–1966 to parents who had recently immigrated to Israel. The interviewees were not chosen at random. I chose them on the basis of four criteria: (a) their parents had immigrated to Israel within 15 years of the declaration of independence; (b) their home language was not Hebrew; (c) their birth place was Israel (to ascertain possible exposure to Hebrew from birth); (d) their birth year was within 20 years of the declaration of independence, while the national identity ethos was prevalent in the country (Lissak, 2009). The study examined the participants' life stories as they reconstructed and reflected on the impact of language on their lives in crucial moments in the history of the Jewish people, the State of Israel, and of the narrators themselves. The participants

Table 1 Participants' profiles

Pseudonym	Parents' country of origin	Parents' immigration year	Year of birth	Languages heard at home
Ben	Iraq	1951	1951	Iraqi Arabic (from parents), Hebrew (from siblings)
Gadi	Poland	After the war	1957	Polish, Yiddish
Noga	Iraq	1934, 1933	1963	Iraqi Arabic in early childhood; then mostly Hebrew
Shaul	Romanian	1949	1951	Romanian, Yiddish, Ladino
Yehuda	Transylvania	1946 (father) 1947 (mother)	1958	Hungarian (occasionally also Yiddish and Romanian when guests visited)

came from different immigration contexts in similar periods. See a summary of participants' features in Table 1.

The interviews took place in neutral places chosen by the interviewees, such as an office at the college, a park, or the researcher's home. They were carried out in Hebrew, as this is the natural default language, the participants' strongest language, and the language of wider communication in Israel. As usual in narrative studies, the interviews were open-ended; the researcher asked clarification or prodding questions but mainly remained silent, letting the narrative take its course. Participants were asked about their childhood experiences with bilingualism in the home vis-à-vis the dominant Hebrew society and their language experiences outside the home. Each interview lasted 50–70 minutes and was recorded and later transcribed by a research assistant. The quotes that appear below were translated by the researcher. Prior to starting the interviews, the researcher informed the participants about privacy and the data protection policy. All the names and places presented in the paper are pseudonyms.

An initial analysis involved reading through the entire data set to get a sense of the contents and themes of the narratives. A thematic coding was then carried out (Miles & Huberman, 1994) after collecting and reading through all the data. To avoid turning narrative analysis into a mere listing of events or themes, Pavlenko and Lantolf (2000) suggest positioning these against an explicit theoretical framework. Accordingly, the themes identified in the narratives are analyzed against the conceptual construct which is the focus of the present volume, namely, *boundary crossing*. Below is a description of the relevant themes and narrative data segments that illuminate the sociolinguistic themes pertaining to boundary crossing.

4 That's the Story: Language as a Boundary Crossing Tool

The narratives collected for this study portray experiences surrounding the shift from the home language brought by parents from the Diaspora to the exclusive use of Hebrew by the children. Inevitably, there is some overlap amongst the themes

presented in different sections, as the home language and the dominant language represent many additional dimensions of one's being. Yet, it is important to start the analysis with a representation of the move across the home/dominant language boundaries.

4.1 Crossing the Boundary Between the Family and the Society

The narratives provided key phrases that highlighted the adults' retrospective reflection on their childhood experiences with the home language and what it represented. Some segments describe a natural transition back and forth across the boundaries and back. Most, however, report on a more drastic movement across boundaries. They moved from seeing the immigrant language as a natural means of communication to perceiving home language use as a divide between the narrators and the Israeliness they saw outside their homes and neighborhoods. Stories of non-negotiable switches to Hebrew-only communication illustrate the children's need to cross the language boundary in order to see themselves as full-fledged, unconditional participants in Israeli society as Israelis.

By abandoning, and often rejecting, the home-language territory and venturing in an all-Hebrew life, they felt closer to their second-, third-, and fourth-generation Israeli peers. Shaul, for example, tells of the rejection of Romanian as an entrance ticket into the wider Israeli society:

*I answered [my parents] in Romanian, I spoke, I mean, everything, life was carried out in Romanian. At around age 10, all of a sudden when...when I felt, I mean, [speaking Hebrew] was the mandate of the youth movement. I started telling [my parents] **'I am not speaking Romanian with you.'***

Yehuda had a similar sudden understanding that in order to be part of Israeli society, he needed to reject Hungarian:

*At around age 10, we used to travel by bus because we didn't have a car, and my mother would speak in Hungarian and I would tell her, "Don't talk to me in Hungarian"... I remember that one time **I gave her an ultimatum**, "If you talk to me in Hungarian again, you will not see me around you." She stopped.*

For Yehuda, this was a milestone in crossing the boundary that life had imposed on him and that prevented, from his perspective, his participation in society as a prototypical Israeli. Through this ultimatum, he walked across the home/outside language boundary to a new identity, one that had no room for the home language. Noga reports on a very similar process of planned and conscious boundary crossing:

***My aversion for the language**, I had an aversion... I tried as hard as I could to distance myself from ... to hide the language. I understood the language, but never said a word in Arabic. Never a word...*

Ben's parents immigrated from Iraq, and the home language was Arabic. Ben continued speaking Arabic with his parents until they died. Today, he uses Arabic at work. As a child, he lived in a mixed Jewish-Arab town and his playmates were from Arab families; the language of communication amongst them was Arabic. This apparently blurred somewhat the boundary between home language and the language of wider communication in the context of a mixed town. In Ben's narrative there are no elements of rejection of the home language in favor of the societal language. In his case, the boundary crossing was less obvious, and perhaps less necessary than in other cases. Ben does not speak of identity conflicts or of definite divides between the home and the outside:

[At home] the dominant language was Arabic. In the outside environment I spoke Hebrew, but the move from one to the other was not problematic, no, it flowed well.

By contrast, Gadi tells us of two full and self-standing worlds in which he felt completely comfortable; he felt no urge or need to cross over from one side to the other. His experience with boundaries between outside-language *territory* and home-language *territory* was different from the experiences of the other children:

[Switching from Hebrew to Polish] was rather automatic. Look, the door, the door would open at home and this was Polish territory.

His house was teeming with Polish: books, newspapers, neighbors. He speaks of Polish-language artifacts with positive feelings:

*My parents would order Polish books by mail, and **when they arrived there was a ceremony**. They would come tied in a string and my mother would cut it. She wouldn't open the package. It was forbidden to see what's inside. He [my father] would open it, of course. He was the man and he would take the books out first.*

Gadi mentions the importance of not having crossed the boundary between the home language and the outside language. For him, the preservation of the home territory seemed to have enormous significance; for example, a filial-parental connection was enacted by maintaining the Polish language. The interplay amongst the language, affect, and affection can be seen in many segments of this narrative:

*My connection with my parents was by means of this language. I mean, we must understand that their world was completely destroyed. A great part of their families was murdered [in the Holocaust]. The only thing that they salvaged from there was the connection with the culture and the language; so I felt, I think that pretty early I understood that **by speaking the language, I can preserve something** from their maps of consciousness.*

When his parents separated and his father later built a new, Hebrew-speaking family, Gadi made use of Polish as a way of creating an uncrossable boundary between his previous family and his father's new one:

*I didn't want them to understand...we spoke, I mean, **there was intimacy in our conversations that distinguished me as his son**, his only son, his eldest, in the context of my sister, the adopted daughter, and his wife.*

Again stressing the preservation of knowledge and closeness through language, Gadi tells us that *not* crossing boundaries helped maintain strong bonds with his

family. With a sense of responsibility, Gadi accepted and respected the salient home-outside language boundaries. When speaking of modern-day immigrants, Gadi mentions grown children who cannot comprehend certain life experiences of their parents. He shares a different experience:

With us there is no such colossal divide. When my mother recites in her old age poems in Polish...so, that's the way it is, I mean, it is a kind of intimacy that is (was) preserved.

For Gadi, then, the boundaries of the home language represented a positive emotional space of affection and intimacy. Although a distinct boundary existed between the home language and the outside language (*'The moment the door opened, it was Polish'*) he navigated smoothly the movement from one context to the other:

*My heart goes out to speak with them the language that will be touching for them. Also today, if I want to make my mother laugh or to say something moving I will say something or recite something in Polish because in Hebrew it won't work. **There is a very sensual layer of the language.***

Interestingly, Gadi also mentions the broader boundary between the (then-immigrant) Northern Tel-Aviv neighborhood and Central Tel-Aviv:

*On the streets you could hear Polish and Yiddish, so, all that "Hebrew, speak Hebrew" thing –that's nonsense, what are you rambling about? Maybe in Dizengoff street, **but here in Ramat Aviv – who spoke Hebrew?***

Continuing the description of his neighborhood, Gadi depicts a remarkable, graceful language movement within and between home boundaries that characterized young Israel's melting pot:

*My building was a Babel Tower of languages, with Bukharians, Yemenites, Persians, it was very natural...people come from different places and all is mixed...in one apartment they spoke Iraqi, next door Persian, etc. etc. I didn't understand it but there was some beauty to it, yes?, that you have other sounds, that there is an enormous polyphony...from everywhere a different sound and people put on a record so each one with his heritage. It seems very beautiful to me. **From a young age I liked this heterogeneity.***

From the narrative segments above we get a taste of the two worlds and identities that the children of immigrants found themselves negotiating in the historical period of post-Holocaust, post-independence Israel. From an analysis of these childhood experiences we can see that similar bilingual upbringings can lead to diverse reactions on the part of the child regarding the mediation between languages and what they represent.

The perception of the divide between the home language and the language(s) outside the home was replete with additional issues besides communication and ethnic identity. The section below offers us a deepened understanding of the relationship between language and identity in Israel in the 1960s and 1970s in the context of the boundaries established between the home/dominant languages.

4.2 *Crossing the Boundary from a Diasporic Identity to an Israeli Identity*

A multitude of personal and social challenges accompanied post-Holocaust immigration to Israel. Providing dwelling to all immigrants was complex, and often neighborhoods became all-immigrant areas. Families had parents from different countries (in this study, representatives from Romania, Hungary, and Iraq), different cultures (European vs. Middle-Eastern, a.k.a. *Mizrahim*, Hebrew for 'Oriental'), and different life experiences (Holocaust survivors vs. Jews who fled before the Holocaust or who came from countries not impacted by it). For children being raised in these neighborhoods, the Israeli experience and way of life was often outside of the scope of their childhoods. Children grew up speaking languages other than Hebrew, and their upbringing resembled that of the countries their parents had come from. To become authentic Israelis, children had to cross the boundary that separated the Diasporic and the Israeli *modus vivendi*. The most obvious and tangible way for children to do so was to reject the home language and embrace the dominant language. Narratives collected for this study speak of participants' decision to cross that boundary during childhood. For example, for Shaul this determination was related to an event with a peer whom he considered a "real Israeli":

Listen, there was an experience, one of the hardest ones for me, till this day. I remember I was on the train going to the Holocaust Day Memorial with the youth movement. Some of the youth were children of immigrants and some were Sabras² (stressed word), originals, first, second, third generation, I don't know. At some point I said, 'my knees hurt'³ ... And then I remember a girl with long hair wearing shorts, a Sabra, a kibbutznik, who said "the correct word is coavOT, not coavIM". I was so terribly offended...oh I failed in the local language....and then I said, I will show her (metaphorically to all children of Israeli-born parents) how one day I will teach their children Hebrew, and I will speak and write, and write poetry...as a journey, a journey to revenge this offense...

Following this event, it became evident to Shaul that the societal language was the ultimate tool needed to become an Israeli and relinquish the familial presence of the Diaspora:

I started to conquer the language. First I said, 'I must conquer the language,' I conquered it...and this distanced me from my native language, the language of the [home] culture.

For Noga, who came from an affluent Iraqi family, Hebrew gradually infiltrated the home and by the time she was in elementary school her parents had stopped speaking Arabic, even between themselves. Her story tells us of the intricate interplay amongst language, culture, and ethnicity. Iraqi immigrants, even if affluent, were initially perceived by Western-European immigrants as having a lower status. As such, Noga tried to hide as much as possible her ethnic background by rejecting the language and culture of Iraq. Living in a mostly Ashkenazi neighborhood, Noga

²The *Israeli Sabra* is the denomination of the native-born Israeli.

³In Hebrew, adjectives and verbs are gendered. Shaul used the wrong grammatical gender in this sentence.

was “probably the only Iraqi child in school.” By the age of seven she had stopped speaking Arabic altogether, apparently as part of the overall rejection of the Iraqi culture. She turned her back on Arabic as she confidently walked across the boundary that separated between the Diasporic perception of the self and the Israeliness she aspired to embrace:

*I perfectly remember [that I said to myself] ‘If I speak Arabic, **I will not be Israeli.** My cousins [who spoke Arabic] will remain immigrants.’*

Also from an Iraqi background, Ben felt that by hiding his origins, he would not be looked down on by Israelis of Western-European background. He had instrumental motives to cross the boundary from an Arabic-speaking Iraqi Jewish Israeli to a melting-pot unlabeled Israeli:

*I wanted to **hide my ethnic background.** My background includes language, culture, everything together. I did it not out of shame, but out of fear that maybe [army officers] won’t behave to me nicely, won’t give me a good service position or who knows what.*

Besides this retrospective reflection on the process she went through as a child, Noga sees herself today as an entity somewhat detached from the labels that often accompany immigrants:

*The way I have developed, what I am, has **no traces** of my being from Iraqi descent.*

The seemingly arduous mission of becoming exemplary Israelis and balancing two identities resulted in boundary-crossing actions also taken by Shaul. He had stopped speaking Romanian and had chosen a life path in which the language of the Diaspora was not present. In this poignant testimony, belonging to the Diaspora represented a burden and a handicap:

*It was clear to me at age 12, 13, 14 that I am Israeli, Hebrew, Zionist, I will build the kibbutz, I will go to the army, I will do everything that “Working Zionism” dictates. It was already clear to me at the age of 11, 12 that I am doing things as part of my separation from **the hump of the Diasporic identity**...I saw the Hebrew language as having a considerable part of my new identity.*

Yehuda, in turn, a child of Holocaust survivors, makes constant and strong links between Hungarian – the language of the Diaspora spoken by his parents – and the image of the Holocaust:

*I had no problem [speaking Hungarian] at home. Outside, I was ashamed... A language from ‘there’...a language of lambs who were murdered, I was not proud of it. This language that stole us, a whole humanity. **I don’t want to be similar to them.** It’s a rejection.*

The following example from Yehuda’s narrative represents a touching and bluntly expressed will to cross the boundary between the past and the present. He recounts how he consciously and decisively moved from one territory to another, from being a youth in the shade of the Holocaust to being a strong, healthy, modern, new-generation Israeli:

*I didn’t want to be part of it...there is nothing to imitate here, nothing to be bonded to...I want ‘that’ experience, **I want the next experience.***

When asked whether this move to the “next” episode included the Hungarian language, Yehuda answered:

It includes it. Anything that is a marker, I don't want it. I want to shrug it off of me...they didn't come with...the Portuguese language. They came with the language of 6 million on skewers [reference to the 6 million Jews murdered in the Holocaust]. As a child you can't understand how they got to that situation, you don't want to understand, you're a child, I want here and now. [I want] respect.

4.3 *Crossing the Boundary from Shame to Pride*

The narrative segments in this section present how the narrators enacted the crossing from shame to pride by means of language, albeit in distinct ways at times. Some narrators, already as adults, made their way from being ashamed of the home language to being proud of it. For others, it was the rejection of the home language as children, and the complete loyalty to Hebrew that helped them cross the boundary from shame (of the home language) to pride (of being Israeli).

For Yehuda, for example, the boundary crossing experience from Diasporic Jewishness to Israeli Jewishness by means of language choice reflects a tight relation between Hebrew and pride. He became deeply ashamed, as a child, of his Hungarian-speaking parents' Holocaust experiences, subsequently finding in Hebrew a firm national identity marker. What Hebrew represented came in stark contrast with what Hungarian represented: novelty, independence, pride, strength, and victory. Interestingly, Yehuda names the language “Israeli” and not Hebrew (for a discussion on the Hebrew language vs. the so-called “Israeli” language, see Zuckermann, 2008). This stresses, even 50 years later, that the language is the symbol of Israeliness, and language choice is deeply associated with national identity:

The Israeli language. This means pride. It's standing tall. All the rest are not languages, I did not appreciate them, including Arabic. During the Six-Day War I was 9 years old, full of pride, full of pride...You grow up with the feeling that “who wants to hear these languages, these ‘loser languages’⁴”, I don't want to hear them – not Berber (Moroccan Arabic), not Yiddish, not Hungarian, not Romanian... Let's get rid of all of those. We have our own pride. We are the Israeli winners. Not any other language, I don't want them.

The shame-pride divide provided also a terrain for different types of crossing. Ben, for example, reports on a move from shame to a peaceful relationship with the language of his childhood:

At a certain age, at one period in my life, I was ashamed of Arabic, but today, I don't want to say “I'm proud of it,” I think it's a silly sentence, but I completely live in peace with it.

Having put Arabic in a separate, tight compartment, it took Noga a few decades to cross back over the boundary from the territory of shame to the territory of

⁴In the Hebrew original, based on the sentence structure, it was clear that “loser” was an adjective to describe the languages and not their speakers.

appreciation. As a child, she was ashamed of her parents and their culture, their language, their accent:

Look, I think I was a bit ashamed of my parents, especially my mother, she was kind of more Oriental. She had a stronger accent...she worked with other Iraqis so she remained [in the original culture] a bit. So... ummm...I was ashamed. I felt uncomfortable bringing friends home.

It is noticeable that by rejecting her heritage language, Noga built an uncontested divide between the two cultures she was immersed in: the home culture she tried to hide and the outside culture she strived to embrace. A turning point for Noga occurred when she watched a movie that dramatized the Jewish Iraqis' life in Iraq. Much to her surprise, Noga's experience with this sudden exposure to the language was a positive one, and through the renewed encounter with the language, Noga crossed the boundary from shame to pride:

*There was the movie...that was 4 years ago. I was in shock, because the movie was in Arabic and I understood it, and it is like it came back to me, because all that childhood experience returned. **But this time I was not antagonistic, I was really happy. Happy that I have that inside me, happy that I've come a long way, and I am also really proud, proud of the Iraqis...** I saw that, and suddenly it is as if this became connected with a willingness to embrace this identity, and not reject it. It was that movie. I thought I didn't know Arabic and all of a sudden I told myself, "I know Arabic!"*

Following this initial reconciliation with Arabic, language became the very same tool that Noga utilized, decades later, to cross, in the opposite direction, the boundary she had built between shame and pride. Beyond that personal boundary, Noga saw her knowledge of Arabic as a tool to bridge the boundary between Jews and Arabs in Israel:

*Now there are more [Arabic-speaking] students, and I open myself to the language...and I really want to know the other. There is no better co-existence than understanding the other's point of view, and I understand this is done by means of language. I am very happy to see now that it [the language] is a part of me. With the Arabic-speaking students, I see we have a common basis, and the **language really removes fences.***

For Shaul, this juncture in his life happened when his parents died. Similarly to Noga and Yehuda, he had stopped speaking the home language as a child. Just as Yehuda, he had given his parents the warning that he should not be addressed in "their" language. Yet decades later, going through his parents' books and music records, Shaul realized he understood Romanian. Already a well-established poet and poetry professor, this renewed encounter with Romanian sparked interest, curiosity, and love for the language:

*I suddenly saw books and music, etc., and I started listening and reading. "Hey, I understand it!" And family and friends came for condolence visits, they spoke in Romanian, and I answered in Romanian. **Suddenly...I felt like it** [felt like speaking Romanian].*

It seems that after having experienced decades of a solid Israeli identity, Shaul was open to a reconnection to a language and culture that seemed rich and intriguing. Just as Noga, Shaul found himself crossing the boundary between shame and pride

again, this time in the opposite direction. Once more, language was the tool to cross that boundary, and he immersed himself in the Romanian language and culture:

*I went to a poetry reading and I heard her [the Romanian poet Ana Blandiana reading her poetry] in Tel Aviv... and I simply cried. And I said, wow, **I renounced to such a rich and intense world.***

By “conquering Romanian in its hardest level, poetry,” Shaul was determined to reclaim the property of the language and culture of which he was once ashamed. He started translating prose and poetry by important Romanian poets and writers to Hebrew and carries out research on poetry and prose under the Romanian communist regime. These were steps that materialized his coming-of-age pride regarding his home language. He did that with the clear purpose of getting a better grasp of his parents’ life experience and identity. Shaul crossed the boundary again, this time actively taking with him his bilingual, bicultural identity, not leaving it behind.

In the context of his close bond with Polish as the home language, Gadi did not speak of feelings of shame. He was aware of a boundary, but did not experience boundary crossing in either direction:

*By the way, **I was not ashamed of Polish.**...All that feeling that one must speak Hebrew, speak Hebrew, I didn’t have a problem, there was another language. I also thought from a young age that it’s good that there is another language in the world*

For Gadi, in fact, there was no divide between pride and shame in the context of language boundaries. Pride lay within his own personal boundaries of language knowledge:

*I never had the feeling that I was ashamed of my parents, quite on the contrary, on the contrary, I liked it that I could communicate with them in this language, that they feel natural with me...And my Hebrew was better than theirs so you have **a feeling of pride** that you master a tool that your parents don’t...*

Through his continuous work translating Polish poetry into Hebrew, Gadi received the Golden Cross, a recognition from the Polish president for his work. This is only one of the markers of a link, not a divide, across boundaries, that are visible in his story:

*I felt that through my activities **I was keeping a bridge** [between cultures].*

The *bridge* image adds a refreshing viewpoint to the boundary crossing mentioned by the other narrators. The more fluid co-existence of the two linguistic selves on each side of the perhaps porous boundary between Polish and Hebrew allowed Gadi to walk naturally across the language boundaries established by the bilingualism of his childhood. The renewed encounter with, and the discovery of his love for Romanian gradually built a similar bridge that allowed Shaul to unite the two worlds represented by his two languages. The same applies to Noga, who explicitly mentioned co-existence between Jews and Arabs, fence removing, and perhaps a future bridge.

5 Discussion

5.1 *Mudes*

Four out of the five participants in this study told the story of their conscious decisions to cross the home/environment language boundary. Shaul and Yehuda reported on specific moments in their lives when their language choice decision was made. Noga and Ben did not point to a specific moment or event but did reflect on a crystal-clear realization that they needed to choose Hebrew over their home language if they were to establish a wholesome Israeli identity. These crucial crossroads in a bilingual's life, when pivotal events define how they position themselves in their environment through language, have been named *mudes* (singular: *muda*) (Pujolar et al., 2010).

Mudes are defined as “specific biographical junctures where individuals enact significant changes in their linguistic repertoire” (Pujolar & González, 2013, p. 139). In other words, they refer to changes in bilinguals' lives that occur at a certain point in time. This is usually a short process or change that leads to a meaningful change in the language choices of the bilinguals. This modification carries with it social and personal significance. *Mudes* provide a framework through which we may more easily comprehend how individuals enact their multilinguality in different ways throughout the life cycle, such as starting high school or ascribing to a language ideology (Puigdevall et al., 2018). The stories shared in the present study, however, do not point to chronological reference points, such as starting primary school, secondary school, or the university; when starting to work; marrying; or when becoming a parent (González et al., 2009). Rather, the *mudes* here stemmed from events that were related to language ideology and ethnolinguistic/national identity.

The notion of “biographical junctures” applies directly to the conspicuous turning points after which a radical language choice was made, namely, relinquishment of the home language. For example, Shaul, Yehuda, and Noga crossed over the thick boundary of the home language to Hebrew and never looked back on the deserted terrain again – only to cross back over that boundary decades later. Their ideology-based *mudes* assist us in making sense of their life trajectories regarding their linguistic repertoire. Ben's *muda* can be identified in his narrative but emanates a more flexible nature. He crossed the language boundary to avoid stigmas but knew how, and chose to, navigate the constant motion between the home and the outside, establishing specific contexts for Arabic/Hebrew language use. Similar bilingual co-existence in oneself (albeit in completely distinct circumstances) has been previously reported by Zentella (1997).

For Gadi, in turn, “biographical junctures” served to essentially *strengthen* his familial connection, intimacy, and bonding through continual use of Polish. For him, there were no *mudes*, as the realization of the impact of language choice did not lead to an alteration in language practice. Rather, such realization brought about a robust ongoing bilingual and bicultural experience which included a sense of

responsibility towards his parents and the will to rectify, as little or as much as he could, the uprooting that followed the most horrific events they had experienced.

5.2 *Home, Shame, and Identity*

The relationship between language and identity has been widely discussed in sociolinguistic research, often focusing on immigrants and their children. In fact, it is during adolescence that the fundamentals for the development of one's identity are set (Phinney et al., 2001). These two forces (language and identity) may cause an internal conflict within children of immigrants and may be manifested as shame. Instances of children ashamed of their parents' language have been observed in the studies about immigration and language; for example, Bonner (2001) discusses situations very similar to the ones experienced by some of our narrators. It seems that the home language carries with it such critical significance for the formation of the self that children in the process of shaping their own identities may find it challenging to adopt a bilingual/bicultural identity. Children in Machowska-Kosciak's (2012) study indeed faced some internal conflict in this regard.

The stories in the present study point to a period during which youth explored and defined their own identities while considering their connection with the home language. This period in their life cycles coincides with what Xavier Vila et al. (2020) identified as one that "speed[s] up the (re)construction of [youths'] personalities" (p. 1160). Similarly, Bonner's study illuminates a point in Garinagu youths' life cycles when it became unquestionable that they had to make a decision between "indicating their adherence to an ethnic identity [by speaking Garifuna, the home language and] ... indicating their status as *authentic Belizean citizens*" (p. 85, emphasis mine). This authenticity was concretized by the use of the dominant language, English Creole; the same authenticity achieved by speaking Hebrew amongst the narrators in the present study.

There are several ways and contexts in which language and national identity are intertwined (for a typology see Szul, 2009). In Israel's early years, language helped materialize the establishment of the country, hence the utmost importance of Hebrew use for youths' sense of belonging to the country as *authentic citizens*. The narrators explain how the home language represented an identity that corresponded with memories or ideals that the new generation depicted in this study chose to preserve or to reject. The use of Yiddish, for example, a language that Yehuda refused to speak or even respond to, has been identified as a "constant reminder of an oppressed Diaspora condition" (Safran, 1992, p. 406), precisely the reason he mentioned for rejecting both Hungarian and Yiddish. It seems that some of the participants ascribed to the new national identity to escape shame.

The narratives discussed in this chapter exemplify the language and identity choices made by our narrators. More importantly, not only do the narratives depict the surface-level family language policy adopted, but they also allow the reader to delve into the intricate depths of the personal and emotional processes Shaul,

Yehuda, Noga, Ben, and Gadi underwent in formative years of their youth and at later stages of their life cycle. Their poignant stories uncover how these processes occurred while, as youth, they were navigating boundaries between language and identity, whether staying within them, crossing them once, crossing them twice, or keeping a fluid movement across those boundaries. The use of the societal language by participants seemed to generally indicate a membership to the national identity that was the crux of the Israeli ethos of the time (Lissak, 2009). In this respect, Bucholtz and Hall (2005) see identity as "something that is constituted through social action, and especially through language" (p. 588). Using Hebrew seemed to be the social action in which our narrators engaged, reflecting the ideology of the unique circumstances of that state-building era.

Edwards (2009) points to a central factor to consider when discussing ethnic identity – "the amount of importance to be accorded to group boundaries" (p. 157). Although not necessarily connected to *ethnic* identities, the narratives studied here reveal the importance given to group boundaries. Thus, the boundaries between the *home-language* group (parents, extended family, adult neighbors) and the *language-of-the-environment* group (youth movement, school, neighborhood children) had an impact on the protagonists of the stories collected for this study. The stories elucidate that there is a close connection between boundary crossing and identity formation in light of a decisive period in the history of the Jewish people and in the establishment of a "modern Jewish identity" (Spolsky, 2013). Spolsky also notes that "Jewish emancipation in modern times posed new choices both of identity and language" (p. 183). The uncontested impact of this momentous period on the boundary-crossing choices of the narrators emerges through the stories they told. The link they draw between language and the formation of the self concurs with the viewpoints that language is a "powerful pillar of identity" (Spolsky, 2013, p. 205) and is used to delineate and enact different identities (Doran, 2004; Giles & Byrne, 1982).

In this context, it is paramount to refer to the revitalization of Hebrew when discussing the dialogic (Bakhtin, 1986) (re)structuring of personal identities of the participants in the present study. The rebirth of Hebrew as a spoken language (as opposed to a liturgic language) was a key factor in the development of a modern Jewish national identity, alongside that of an Israeli identity. Indeed, Spolsky (2013) sustains that one of the factors contributing to the success of the revival of Hebrew is the solid ideology of Zionist movements that worked towards the return of Jews to the land of Israel. This view is supported by Edwards's (2009) perception of language as an essential component of nationalism.

The narratives presented here provide instances of this view. We learn from the stories that mastery of Hebrew as a native language and total dedication to it, at the expense of knowledge and use of the home language, went much beyond the family language practice realm or immigrant children's choice of identification with the language of the environment. Language choice here served as a declaration of group membership to the nation and ascription to a national identity. Accordingly, Kuzar (2001) stresses that "this strong sense of belonging to a collective identity... entails social discipline and self-policing, as well as self-sacrificing practices..." (p. 134).

Some of our narrators sacrificed part of their family legacies when establishing an uncontested membership to the national identity. Others lived at peace on both sides of the language boundaries. This may be attributed to the way they *interpreted* (Li, 2000) the social change of which they were active participants: the formation of a modern Israeli identity through the use of modern Hebrew in daily life.

Being a Hebrew speaker, whether monolingual or bilingual, became the pulse and the breath of the personal and collective identity which dominated the discourse and the ideology of children of immigrants building the new State and a modern national identity. In this regard, it has been stated that emotional bonds established with a collective ethos have a potential impact on individuals' way of thinking (Bartal et al., 1998; Dahan-Kalev, 2005) and, by extension, on the language choices they make. The different ways in which the participants in this study chose to cross the boundaries of the home language reflect to a great extent the identity they aimed to embrace in the context of the Israeli society of the period.

5.3 Boundary Crossing in Language Education

In an era of global migration, many schools face challenges in the education of bilingual children. Bi/multilingual children and their families face the challenges of societal-language schooling, of heritage language maintenance, and multilingual intergenerational communication. It seems that narrative research, through the scrutiny of life stories, may assist researchers in grasping more effectively the process that all parties involved go through – children, families, classmates, and educators. This personalized inquiry may serve us in defining language policies suitable in each specific case.

Through the narratives presented and analyzed in this chapter, we are able to acknowledge that boundary crossing by bilingual children might contribute to language education and language education research. First, it can assist educators and social workers in appreciating the plight of first- or second-generation immigrant. The complexity of the double purpose of strengthening the self while promoting integration (Rosén & Bagga-Gupta, 2013) can be enlightened by careful inspection of bilinguals' narrated childhood experiences. Often focused on linguistic abilities, language educators may be less aware of the weight of the navigation across boundaries that bilinguals endure, thrive on, aim at, or evade from – to name just a few possibilities in a wide range of experiences related to language learning contexts.

Moreover, this narrative study indicates the need to employ a multitude of methodologies in bilingualism research to empower language learners and language educators. The personal narratives reported in this chapter assist us in achieving a deepened knowledge of language learning and emotional processes undergone by bilingual children in immigrant contexts. This insight is an invitation for language educators to elicit stories of each bilingual child and be attentive to their experiences navigating the boundary between the home and the society. This boundary crossing

seems to have a significant impact on the bilingual self, indirectly acting upon language learning and education.

6 Conclusion

The analysis of the narratives in this chapter showed that, overall, individuals saw in language choice an enhanced sense of responsibility towards themselves and their families. Limitations should be acknowledged, indicating possible areas for future research. First, the participants all grew up in the city. It would be interesting to see if children growing up on a kibbutz had a different experience than the ones reported on here. Also, the immediate environment in which they grew up was diverse: an affluent town, a mixed city, immigrant neighborhoods. These have a direct impact on the children's social milieu in school and with neighborhood friends. It can be useful to analyze narratives that narrow down on the children's experiences outside the home, according to the social environment in which they were immersed. In addition, the age range of the participants (15 years) may have been too wide, since in those crucial years of the beginning of the country, social changes occurred (and still do so) at an extremely fast pace, and the narrators' age-related/period-related realities may have been different. Research on individuals of the same age group may yield a different picture of the movements through boundaries in the sociolinguistic context of this study. Finally, due to length limitations, the theme of emotion could not be developed further. It will be interesting to explore emotional bonds with native bilinguals' languages and the emotional significance of keeping a home language or returning to a relinquished one.

This chapter investigated one aspect of the linguistic trajectories bilingual children growing up in Israel following events of tremendous proportions for the Jewish people. Both the Holocaust and the establishment of the State of Israel impacted deeply the language choices and the attendant childhood experiences of the adults who shared their stories for the present research. Through their narratives, they reconstructed and reflected on the impact of language on their lives. Their life stories portray the navigation among boundaries of the self, the home, and the collective as a journey with specific turning points and key events. They uncover emotion-loaded moments and perceptions of the symbolic meaning of the societal language, the home language, and the boundary between them. We learn from our narrators the intensity of actions and sentiments relating to boundary crossing of all sorts; we also learn that this crossing can occur in different directions across the lifecycle. Above all, we see that regardless of how the language boundaries were negotiated, it took an immense amount of courage to cross them one way, to cross “there” and back, or to live peacefully with those boundaries. Through the stories told in this study we can once more appreciate the complexity of bilingualism and the intricacies of each bilingual individual's relationship with and through language(s) along the course of their lives.

The experiences shared with us in this study potentially assist researchers and educators in interpreting the needs of bilinguals, thus enabling the constructions of culturally and linguistically relevant pedagogy. Such a pedagogy can perhaps take into consideration familial realities that impact language learning and language engagement, thus facilitating a mediation of identities and of linguistic skills by immigrants and children of immigrants.

Acknowledgement I thank the participants in this research, “Shaul,” “Gadi,” “Noga,” “Ben,” and “Yehuda” for opening their hearts and telling me their fascinating life stories.

References

- Bakhtin, M. (1986). *Speech genres and other late essays*. Trans. Vern W. McGee. University of Texas Press.
- Bar-Tal, D., Jacobson, D., & Kleiman, A. S. (1998). *Security concerns: Insights from the Israeli experience*. JAI Press.
- Bonner, D. M. (2001). Garifuna children's language shame: Ethnic stereotypes, national affiliation, and transnational immigration as factors in language choice in southern Belize. *Language in Society*, 30(1), 81–96.
- Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4–5), 585–614.
- Creese, A., & Blackledge, A. (2015). Translanguaging and identity in educational settings. *Annual Review of Applied Linguistics*, 35, 20–35.
- Dahan-Kalev, H. (2005). “Generals” in education: The military as principals. In H. Gur (Ed.), *Israel affairs (in Hebrew)* (pp. 109–129). Yedioth Aharonot and Hemed Books.
- De Fina, A. (2015). Narrative and identities. In A. Georgakopoulou & A. De Fina (Eds.), *Handbook of narrative analysis* (pp. 351–368). Wiley-Blackwell.
- Doran, M. (2004). Negotiating between Bourge and Racaille: Verlan as youth identity practice in suburban Paris. In A. Pavlenko & A. Blackledge (Eds.), *Negotiation of identities in multilingual contexts* (pp. 93–124). Multilingual Matters.
- Dubiner, D. (2018). ‘We don’t think about it, we just mix’: Language choice and ethnolinguistic identity among Arabic-Hebrew bilinguals in Israel. *International Journal of Bilingual Education and Bilingualism*, 24(2), 191–206.
- Edwards, J. (2009). *Language and identity: An introduction*. Cambridge University Press.
- First, A., & Avraham, E. (2010). Contesting national identity during crisis: The use of patriotism in Israeli advertisements. *Communication, Culture & Critique*, 3(3), 334–351.
- Fishman, J. A. (2000). Language planning for the “other Jewish languages” in Israel: An agenda for the beginning of the 21st century. *Language Problems and Language Planning*, 24(3), 215–231.
- Fishman, J. A. (2013). Language maintenance, language shift, and reversing language shift. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism* (pp. 466–494). Wiley-Blackwell.
- Giles, H., & Byrne, J. (1982). An intergroup approach to second language acquisition. *Journal of Multilingual and Multicultural Development*, 3(1), 17–41.
- González, I., Pujolar, J., Font, A., & Martínez, R. (2009). *Entre la Identitat i el Pragmatisme Lingüístic. Usos i Percepcions Lingüístiques dels Joves Catalans a Principis de Segle*. Generalitat de Catalunya.

- Gu, M. (2011). Language choice and identity construction in peer interactions: Insights from a multilingual university in Hong Kong. *Journal of Multilingual and Multicultural Development*, 32(1), 17–31.
- Hall, C. J., Smith, P. H., & Wicaksono, R. (2017). *Mapping applied linguistics: A guide for students and practitioners* (2nd ed.). Routledge.
- Hamers, J. F., & Blanc, M. (2000). *Bilinguality and bilingualism*. Cambridge University Press.
- Kuzar, R. (2001). *Hebrew and Zionism: A discourse analytic cultural study*. De Gruyter.
- Leeman, J., Rabin, L., & Roman-Mendoza, E. (2011). Identity and activism in heritage language education. *The Modern Language Journal*, 95(4), 481–495.
- Li, W. (2000). Towards a critical evaluation of language maintenance and language shift. *Sociolinguistica*, 14, 142–147.
- Lissak, M. (2009). *Studies in Israeli social history (in Hebrew)*. The Bialik Institute.
- Machowska-Kosciak, M. (2012). A language socialization perspective on knowledge and identity construction in Irish post-primary education. In F. Farr & M. Moriarty (Eds.), *Language learning and teaching: Irish research perspectives* (pp. 87–110). Peter Lang.
- Menard-Warwick, J. (2004). I always had the desire to progress a little: Gendered narratives of immigrant language learners. *Journal of Language, Identity, and Education*, 3(4), 295–311.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.
- Mu, G. M. (2015). A meta-analysis of the correlation between heritage language and ethnic identity. *Journal of Multilingual and Multicultural Development*, 36(3), 239–254.
- Myers-Scotton, C. (2005). *Multiple voices: An introduction to bilingualism*. Wiley-Blackwell.
- Neuman, S. (1999). *Aliyah to Israel: Immigration under conditions of adversity*. Retrieved from www.econstor.eu/bitstream/104https://19/20923/1/dp89.pdf
- Pavlenko, A. (2007). Autobiographic narratives as data in applied linguistics. *Applied Linguistics*, 28(2), 163–188.
- Pavlenko, A., & Blackledge, A. (2004). *Negotiation of identities in multilingual contexts*. Multilingual Matters.
- Pavlenko, A., & Lantolf, J. (2000). Second language learning as participation and the (re)construction of selves. In J. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 155–177). Oxford University Press.
- Phinney, J. S., Romero, I., Nava, M., & Huang, D. (2001). The role of language, parents, and peers in ethnic identity among adolescents in immigrant families. *Journal of Youth and Adolescence*, 30(2), 135–153.
- Puigdevall, M., Walsh, J., Amorrortu, E., & Ortega, A. (2018). ‘I’ll be one of them’: Linguistic *mudes* and new speakers in three minority language contexts. *Journal of Multilingual and Multicultural Development*, 39(5), 445–457.
- Pujolar, J., & González, I. (2013). Linguistic ‘mudes’ and the de-ethnicization of language choice in Catalonia. *International Journal of Bilingual Education and Bilingualism*, 16(2), 138–152.
- Pujolar, J., González, I., & Martínez, R. (2010). Les mudes lingüístiques dels joves catalans. *Llengua i ús: revista tècnica de política lingüística*, 65–75.
- Rosén, J. K., & Bagga-Gupta, S. (2013). Shifting identity positions in the development of language education for immigrants: An analysis of discourses associated with ‘Swedish for immigrants’. *Language, Culture and Curriculum*, 26(1), 68–88.
- Safran, W. (1992). Language, ideology, and state-building: A comparison of policies in France, Israel, and the Soviet Union. *International Political Science Review*, 13(4), 397–414.
- Seviñç, Y. (2016). Language maintenance and shift under pressure: Three generations of the Turkish immigrant community in the Netherlands. *International Journal of the Sociology of Language*, 242, 81–117.
- Smootha, S. (2008). The mass immigrations to Israel: A comparison of the failure of the Mizrahi immigrants of the 1950s with the success of the Russian immigrants of the 1990s. *Journal of Israeli History*, 27(1), 1–27.

- Spolsky, B. (2013). Hebrew and Israeli identity. In Y. Suleiman (Ed.), *Language and identity in the Middle East and North Africa* (pp. 189–200). Routledge.
- Spolsky, B., & Shohamy, E. G. (1999). *The languages of Israel: Policy, ideology, and practice*. Multilingual Matters.
- Szul, R. (2009). Language, nation, and state: Past and present. *Economía, Sociedad y Territorio*, IX(31), 561–579.
- Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 3–15). Multilingual Matters.
- Wodak, R. (2012). Language, power and identity. *Language Teaching*, 42(2), 215–233.
- Xavier Vila, F., Ubalde, J., Bretxa, V., & Comajoan-Colomé, L. (2020). Changes in language use with peers during adolescence: A longitudinal study in Catalonia. *International Journal of Bilingual Education and Bilingualism*, 23(9), 1158–1173.
- Zentella, A. C. (1997). *Growing up bilingual: Puerto Rican children in New York*. Blackwell Publishers.
- Zuckermann, G. (2008). *Israelit safá yafá (Israeli, a beautiful language: Hebrew as myth)*. Am Oved.

Boundary Crossing from the Start: 55 Years of Second Language Grammatical Gender Research in Review



Daniel R. Walter

Abstract In this chapter, I explore how the study of grammatical gender has been a prime example of the impact that boundary crossing can have on the understanding of a phenomenon by creating nuance through the incorporation and juxtaposition of multiple perspectives. In the first section, I start with a description of grammatical gender and its different instantiations (or absence) across languages, including cue patterns, agreement structures, and noun-class parallels. I then move on to the ways in which grammatical gender has been investigated and how this research has crossed theoretical and methodological boundaries. The third section focuses on grammatical gender in SLA and aims to deconstruct the boundary between theory and practice, with a focus on what functional approaches have taught us about instruction of grammatical gender. The fourth section focuses on the history of grammatical gender instruction and the various pedagogies that have been tested to support the teaching of grammatical gender, including innovative approaches based on functional and sociocultural approaches. In the final section, I reflect on the many boundaries crossed: theoretical, methodological, disciplinary, linguistic, developmental, typological and, maybe most importantly, temporal. I end with a discussion of the future boundaries to be crossed in this area of research.

Keywords SLA · Grammatical gender · History · Boundary crossing · Teaching

1 Introduction

In 1967, G. Richard Tucker completed his thesis for his PhD from McGill University under the title *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. This thesis focused on first language (L1) French speakers

D. R. Walter (✉)

Humanities Division, Emory University, Oxford College, Oxford, GA, USA

e-mail: daniel.walter@emory.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching,
Understanding, and Improving Language Education*, Educational Linguistics
58, https://doi.org/10.1007/978-3-031-24078-2_3

and the rules that govern French grammatical gender marking. In his own words, Tucker (1967, p. 38) explains that “the study of gender may prove to be a relatively easily explicable example of rule acquisition in language, and the process of attempting to explain this acquisition may provide a useful model for more complex linguistic features (...)” (p. 1). One year later, Tucker published an article with colleagues Lambert, Rigault, and Segalowitz titled “A psychological investigation of French speakers’ skill with grammatical gender.” Here, Tucker uses a different vocabulary to talk about French grammatical gender:

In the case of first-language learning, it seems apparent that the French child is able to distinguish and utilize, by induction, from the recurring regularities in the language, those patterns of cues that mark gender. He appears to be very skilled in generalizing from these patterns to novel occurrences. (Tucker et al., 1968, p. 315).

Right from the start, we can see one of the first of many Tuckerian boundary crossings through the vocabulary used to describe a single phenomenon. In the first quote, the words “rule” and “govern” are reminiscent of a Chomskian perspective on language acquisition. In the second, the words “psychological”, “patterns”, “regularities”, and “cues” call to mind a different, still young field of psycholinguistics, which offered alternative ways of describing the nature of language.

This theoretical boundary crossing is not the only one made between these two studies. Other boundaries include the methodologies used to investigate the role of grammatical gender in linguistic systems and their uses, as well as the boundaries between pedagogical approaches. In addition, Tucker’s initial focus on first language acquisition (L1A) of French grammatical gender eventually led him to questions about the role of grammatical gender in second language acquisition (SLA), in addition to studies in languages beyond the borders of French. Over the course of his work, he leaped between typically separate fields to develop a more robust understanding of grammatical gender on the whole, as a phenomenon of study that does not leave itself easily confined within the borders of one language, one discipline, or one theoretical approach.

Over the last 55 years since these two publications, inquiries into grammatical gender in applied linguistics have changed from the Chomskian rule-governed understanding which peaked at the time of the young Dr. Tucker to a data-driven era that has embraced non-rule-based features and incorporated perspectives and methodologies from diverse academic disciplines. Tucker’s early work set the stage for a complex and nuanced understanding of grammatical gender and its acquisition in first and second languages. In this chapter, I explore how the study of grammatical gender has been a prime example of the impact that boundary crossing can have on the understanding of a phenomenon. First, I describe grammatical gender and its different instantiations (or absence) across languages, including cue patterns, agreement structures, and noun-class parallels. I will then move on to the ways in which grammatical gender research has been investigated and how this research has crossed theoretical and methodological boundaries. The third section focuses on grammatical gender in SLA and aims to deconstruct the boundary between theory and practice with a focus on what functional approaches have taught us about instruction of grammatical gender. The fourth section focuses on the history of grammatical gender instruction and the various pedagogies that have been tested to

support the teaching of grammatical gender, including innovative approaches based on functional and sociocultural approaches. In the final section, I reflect on the many boundaries crossed in the research on grammatical gender: theoretical, methodological, disciplinary, linguistic, developmental, typological and, maybe most importantly, temporal. I conclude with thoughts on the future boundaries to be crossed in this area of research.

2 Grammatical Gender across Linguistic Boundaries

The first questions to pose is, what is grammatical gender? It is integral to so many Indo-European languages, but at the same time the term *grammatical gender* fails to properly describe the phenomenon, especially since a similar phenomenon is observed in noun-class systems such as Swahili, which operate in structurally similar ways but differ in semantic overlapping. In fact, in order to really understand the nature and function of grammatical gender, I must cross quite a few linguistic, typological, and semantic boundaries.

2.1 On the Nature of Grammatical Gender

One of the first boundaries we come across when we look at the nature of grammatical gender is the semantic distinction between the terms we use and the objects they refer to. In languages with grammatical gender, all nouns are assigned a grammatical gender on top of and in addition to any naturally occurring semantic gender or sex, in the biological sense (from here on, sex). These grammatical genders and sexes can overlap, but do not have to, and in the grand scope of all nouns in a language, the vast majority of the time the grammatical gender has nothing to do with sex. The gender assigned to a particular noun is mainly a way to capture the similarity among nouns that behave in similar ways. And while the assignment of grammatical genders to nouns is usually aided by phonological, morphological, or semantic similarities among nouns, the designations *masculine* and *feminine* are rather arbitrary.

This previous point about assignment of grammatical gender is particularly important because gender assignment was previously thought to be random in some languages. For example, in German, it wasn't until studies by Zubin and Köpcke (Köpcke & Zubin, 1984, 1996; Zubin & Köpcke, 1981, 1986) that gender assignment was understood to be highly predictable. Rather than a reliance on memorization for all nouns, we now knew that people could rely on cues to assign grammatical gender to the vast majority of nouns. Remaining with German, there are, in fact, phonological, morphological, and semantic cues to grammatical gender. Phonological cues differ from morphological ones in that, for phonological cues, the sound itself is the cue and does not carry meaning. Meanwhile, for morphological cues, the sound is tied to a particular meaning. For example, most words in

German that end in /ə/ like *Katze* [cat] are feminine, but there is no meaning attached to the /ə/ itself; it is simply part of the sound of the word, i.e., the phonological representation of the word, and is, therefore, a phonological cue. On the other hand, the morphological cue /kart/, as in *Glücklichkeit* (happiness) has a morphological function – to turn adjectives into nouns (similar to the function of ‘-ness’ in *happiness*). There are also semantic cues in German. For example, all days of the week are masculine, all alcoholic beverages (except beer) are masculine, and all baby animals are neuter. So within one linguistic system, the cues to gender cut across linguistic categories. These types of grammatical gender cues, as well as their distribution, reliability, and competition with one another cross language boundaries. One language may only have phonological cues to gender, while another only has semantic, and a third has both.

This cross-linguistic difference is also true for the number of grammatical gender categories in a language. Whereas most Romance languages only have two levels, masculine and feminine, that is not true for Romanian, which retained its neuter category from Latin. The dropping of the third gender occurred in all other Romance languages as well as some other Indo-European languages such as the Germanic language Dutch (whose two resulting genders were renamed to *neuter* and *common* in formal linguistic descriptions) and English (which dropped almost all of its grammatical gender over time). This is interesting because closed-class words are less productive and harder to change over time.

2.2 *An Ill-Fitting Name*

Here, I would like to return to the term *gender* itself, because when we discuss typologically different languages, we see similar linguistic features that are not traditionally called *genders* that nevertheless act in very similar ways. So in order to truly understand what grammatical gender is, we again need to cross the boundary between languages or language groups. This effort to cross and examine multiple languages’ and language families’ boundaries leads to a better understanding of the role that semantics and morphology play in languages with grammatical gender.

First, when we compare language with genders to other languages with noun-classes, we can see that much of the confusion arises as a direct result of the ill-fitting term *gender*. This problem in linguistic terminology can be seen in at least four ways:

- the lack of overlap between grammatical gender and sex present in languages,
- the origins of the term *grammatical gender* and an attempt to encase all languages with this typological feature into the same linguistic box,
- the similar function that noun-class systems in non-Indo-European languages play,
- and the possible over-extension of biological sex connotations onto inanimate objects.

First, the overlap of sex with grammatical gender complicates and overemphasizes this relationship. In languages with grammatical gender, all nouns are assigned a gender. In many languages (including German, Spanish, French, Latin, and Greek, among others), sex often overlaps with grammatical gender for people and professions. For example, in German, *die Frau* and *der Mann*, “the woman” and “the man” respectively, agree in both grammatical gender and sex. However, one only needs to look to the word for “girl”, *das Mädchen*, to see the break between grammatical gender and sex. *Das Mädchen* in German (as is evident from the difference in the article-noun agreement pattern visible on the definite article) is not the same gender as *die Frau*. Instead, “girl” in German is neuter. This grammatical gender assignment has nothing to do with the semantic nature of the referent (although, as previously mentioned, other semantic cues for gender assignment do exist in German). Grammatical gender assignment here is related to the morphological marking, specifically the umlaut over the *a* and the suffix *-chen*, which makes nouns diminutive. All diminutive nouns in German are formed through this process,¹ and therefore all diminutive nouns in German are assigned neuter grammatical gender. If not even nouns with semantic cues to sex provide a reliable basis for grammatical gender assignment, then it is not a surprise that in the majority of cases, grammatical gender has nothing to do with sex.

The second problem with the term *gender* is that the origin of this term comes from the study of classical languages. The first known use of the term *gender* to describe noun classes was by the Greek philosopher Protagoras in the fifth century B.C., when he divided Ancient Greek nouns into three classes, *masculine*, *feminine*, and *inanimate* (Aikhenvald, 2004). Greek has a noun-class system typical of Indo-European languages. Over the course of centuries, various philosophers and others interested in the nature of language forced, for better and often for worse, the pre-established categories from Greece onto the grammatical gender systems of other Indo-European languages, seemingly regardless of their fit. Even within Indo-European languages, there can be important distinctions between grammatical gender systems that make this classical three gender system unfit. As previously mentioned, many Romance languages only have two genders.

Finally, the existence of languages with similarly functioning grammatical categories should add another layer of skepticism for using *gender* as the overarching term for this phenomenon. A perfectly reasonable replacement term, *noun class*, exists and can more accurately encompass the grammatical gender systems of all languages as well as other features present in languages like Swahili, which has 16 noun classes and uses categories such as *human*, *animal*, and *plant* as semantic categories to organize nouns into different noun classes.

How much of the linguistic understanding from the study of Greek structures has interfered with our ability to differentiate the functions and forms of grammatical gender systems from the way languages refer to sex? How much has the word *gender* tainted our ability, as researchers and speakers, to understand noun-class

¹Or through the suffix *-lein*, although *-chen* is more productive, especially in modern German.

structures in languages that we observe as having *gender* systems? This question about linguistic relativity has gained increasing scrutiny in recent research as researchers test the boundaries and implications of a weak Sapir-Whorf hypothesis, i.e., the idea that the language(s) we speak do not determine our ways of thinking but do play a role in highlighting certain features of our environment via entrenched grammatical processes (Kay & Kempton, 1984). One of the major claims of a weak Sapir-Whorf hypothesis is that grammar affects what we pay attention to as humans. According to Samuel et al. (2019), “labels or grammatical information hone attention to associated features, which in turn feed back down to lower-level processes in a feedback loop. These effects can be upregulated or downregulated by the salience of the relevant linguistic information in the task,” (p. 1782). Therefore, if a grammar is telling you to pay attention to the way nouns are being categorized by morphologically similar declensional paradigms, then that information is continually being fed back to the processor and assigned a salient role in the input.

When we assign a name like *gender* to these paradigms, which overlap with sex to minor degrees, it is easy to see how a constant influx of gender information about nouns could influence semantic connections. This point is argued by Phillips and Boroditsky (2003), who report that,

A series of studies found effects of grammatical gender on people’s perceptions of similarity between objects and people. This was true even though the tasks were performed in English (a language devoid of grammatical gender), even when the tasks were non-linguistic (e.g., rating similarities between unlabeled pictures), and even while subjects were engaged in a verbal interference task. Finally, results showed that crosslinguistic differences in thought can be produced just by grammatical differences and in the absence of other cultural factors. It is striking that even a fluke of grammar (the nearly arbitrary assignment of a noun to be masculine or feminine) can have an effect on how people think about things in the world. (p. 933)

Studies by Boroditsky (2001) and Boroditsky et al. (2003) seem to provide additional evidence for these effects.

However, these findings are not consistent. In a study by Kousta et al. (2008) that investigated semantic transfer among L1 Italian speakers of L2 English, the authors “found no evidence of transfer from Italian to English of the semantic effects of gender and interpret this lack of transfer as evidence for the constrained role grammatical gender has on bilingual cognition” (p. 854). Similarly, Bender et al. (2011), who conducted a study on L1 German speakers, did not find an effect of language on thought regarding grammatical gender. At this time, it is still unclear how much, if at all, grammatical gender influences a person’s semantic representation of inanimate objects. What is clear, however, is that more research is needed to understand whether grammatical gender crosses the boundary of morphology into semantics.

3 Grammatical Gender Research across Theoretical and Methodological Boundaries

Grammatical gender is a topic of great interest within a number of research perspectives, and although past efforts to describe grammatical gender from these different perspectives have been very fruitful with regard to our understanding of grammatical gender, we only benefit from these multiple perspectives when we allow ourselves to cross the boundaries of various research paradigms and theories.

From formal linguistic perspectives working within the parameters of Universal Grammar (UG), grammatical gender has been a doorway into the study of feature assignment across languages. One of the major questions is where and how grammatical gender information is assigned. In one analysis of grammatical gender in Romance and Bantu languages, Carstens (2010) used grammatical gender to argue that nouns can have “intrinsically valued but uninterpretable” (p. 28) features. On the other hand, Kramer (2014) argued that gender description in Amharic, a Semitic language, relies on both grammatical gender and sex. Rather than separating grammatical gender and sex, the author “developed a gender assignment system that is almost entirely based on sex as an interpretable feature on [nouns], and ‘masculine’ forms as a default for anything that does not have a [+FEM] on [the noun]” (Kramer, 2014, p. 11). Kramer (2014) argued that this reanalysis “is more successful than previous analyses in that all the Amharic facts are accounted for, there is no need for a discourse referent lexicon connection, and there is no more ‘calculation’ of gender from sex” (p. 11). Interestingly, within a UG framework there seems to be an important discussion related to boundary crossing – whether non-syntactic/morphological aspects of language, like semantics, can affect feature assignment.

Beyond formal linguistic approaches, developmental psychology is interested in how children acquire grammatical gender in both L1 and early L2 acquisition. For example, Blom et al. (2008) investigated acquisition of the common and neuter genders on both determiners and adjectives in Dutch among L1 Dutch speaking children, child Moroccan L2 learners of Dutch, and adult Moroccan L2 learners of Dutch. By comparing the three groups, the researchers found that “the vast majority of the children’s errors could be interpreted as use of the common form (i.e., *schwa*-adjective) in neuter contexts and were in this respect consistent with the errors in definite articles,” (Blom et al., 2008, p. 322). From this we learn that child learners’ developmental trajectories may differ from those of adults, and that crossing boundaries between L1A and SLA may lead to further insights about how particular grammatical features develop. In contrast with L1 learners, the adult L2 learners in this study often produced “bare adjectives” (Blom et al., 2008, p. 317); i.e., they did not apply any grammatical gender to attributive adjectives.² This could imply a difference in language acquisition, where children, having fewer defined L1 pathways, automatically process grammatical information in the input, even if the same information is absent in their L1, while adults rely on L1 pathways that would not be

²This result was also found by Walter (2020) for L1 English high school learners of L2 German.

looking for grammatical information in new places (i.e., grammatical gender marking on adjectives), assuming grammatical information is not located in the same linguistic environments in their L1. Unlike L1 learners who are forming an initial pathway to process grammatical gender, L2 learners may be required to either abandon or reform processing pathways to account for gender information. However, the authors point to their results and other studies which indicate that “in bilingual/child L2 development, definite articles do show a (prolonged) development, whereas attributive adjectives seem to fossilize” (Blom et al., 2008, p. 322). Thus, there seems to be a difference in either the relevance or frequency of grammatical gender marking on determiners as opposed to adjectives that is causing a divergence between the acquisitional trajectories of these two parts of speech. The previously described study by Blom et al. (2008) also crossed into the field of corpus linguistics; the source of the data they used to come to their findings was “based on Dutch adult- and child directed speech in three CHILDES corpora” (p. 303).

Researchers within usage-based approaches are interested in similar questions about acquisition as those posed by formal linguists and developmental psychologists, but their methods and theories differ substantially. In a study using an artificial language, Arnon and Ramscar (2012) asked whether the grain size and order of acquisition could affect the acquisition of a novel grammatical gender system. By modifying the grain size (whether there was a large or small boundary between articles and nouns) and the order in which participants saw the relevant grammatical gender structure, the authors showed that participants who saw items with smaller boundaries and/or had the sequence-first condition outperformed other groups on both the forced-choice and production tasks. They argued that these findings “[fit] nicely with usage-based models of language, which posit that grammatical relations emerge from a gradual process of abstraction over stored utterances” (p. 2116). This explanation and these findings contrast strongly with the positions of formal linguists, who have focused on the availability of transfer and access, rather than properties of the input and learning.

Crossing another disciplinary boundary, we find that psycholinguists are also interested in grammatical gender, and in looking at this phenomenon, they collect and interpret knowledge from various mental functions such as acquisition, processing/comprehension, and production. Beginning with acquisition, it is not the case that L1 speakers always pick up grammatical gender so effortlessly, as is so often the sentiment regarding L1 acquisition. For example, cue reliability and frequency have significant effects on the L1 acquisition of grammatical gender (e.g., in French: Matthews, 2010; in German: Mills, 2012). Psycholinguistic researchers are also interested in how the existence of grammatical gender affects processing. One way to investigate this is to measure the time it takes to access gender information using timed lexical decision tasks. Using this method for L1 Spanish number and gender processing, Dominguez et al. (1999) found surface frequency effects for access times by gender, within gender, and between singular- and plural-dominant forms. However, when the researchers compared within-group access times, they found differences between real word items but not between non-word items. This means that semantic information, like connotations between genders and sex, were applied

to real words and slowed down processing but that this additional processing step for semantic information was not carried out for non-words.

So why is there an additional processing step for real words as opposed to non-words? If we cross over into connectionist models, they help shed light on how these structures are processed and the number of routes through which grammatical gender information can be accessed. Studying L1 Hebrew speakers, Gollan and Frost (2001) proposed a connectionist model with multiple routes to gender using a gender decision and grammaticality judgment task. Based on the results, they postulated “a model containing two routes to grammatical gender: one that involves an abstract gender node, and another that is form-based and is assumed to play a greater role in recovery from agreement errors,” (Gollan & Frost, 2001, p. 627). The question of how one “gets” to gender information must then at least raise questions about whether semantic, phonetic, or morphological information is the faster way to access gender and whether these routes to gender are variable based on the way gender is assigned and distributed within a language.

Offline measures can also provide insight into variable routes to gender assignment. Hohlfeld (2006), investigating L1 German speakers’ assignment of gender to non-words, also found evidence for multiple routes to gender within the same language; one they called *lexical* and the other *rule-based*. The author argued that without some semantic referent, as is given with non-words, there must be some other way that German speakers are able to assign gender aside from the learned gender/noun pairing through usage. They argue that “gender assignment might be guided by either gender marking regularities alone or lexical information as well as gender cue information (postlexical checking)” (Hohlfeld, 2006, p. 139). The author points to studies that found similar results (e.g., Bates et al., 1995, 1996; Gollan & Frost, 2001); but they also note that these findings contrast with some on French (e.g., Desrochers & Paivio, 1990; Taft & Meunier, 1998).

The methods that psycholinguists use to investigate grammatical gender go beyond traditional behavioral ones. In recent years, researchers have begun to use functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) to look at the brain during processing. For example, de Resende et al. (2019) used EEG to look at event-related potentials (ERPs) of distinct neurocognitive mechanisms for different grammatical gender conditions. In this study, L1 Brazilian Portuguese speakers “read sentences containing congruent and incongruent grammatical gender agreement between a determiner and a regular or an irregular form (condition 1) and between a regular or an irregular form and an adjective (condition 2),” (de Resende et al., 2019, p. 181). After analysis of the ERP results, the researchers found a “LAN/P600 effect for gender agreement violation involving regular and irregular forms in both conditions,” (de Resende et al., 2019, p. 181). The authors argued that these results “suggest that gender agreement between determiner and nouns recruits the same neurocognitive mechanisms regardless of the nouns’ form and that, depending on the grammatical class of the words involved in gender agreement, differences in ERP signals can emerge,” (de Resende et al., 2019, p. 181). This ERP effect on grammatical gender non-congruence has been fairly well established and can have variable effects based on the way the grammatical gender

system is distributed within a language. In another example, Caffarra et al. (2015) found that among L1 Italian speakers, transparent nouns, i.e., ones that had clear and reliable phonological cues to gender, “elicited an increased frontal negativity and a late posterior positivity compared to irregular nouns (350–950 ms), suggesting that the system is sensitive to gender-to-ending consistency from relatively early stages of processing” (p. 1019). One possible interpretation of this data might be that while one route, either lexical or morpho-phonological, might be faster in terms of access, the strength of the connections might be increased when both routes are available.

Psycholinguists can also cross methodological boundaries within the same study to provide a more detailed, triangulated explanation of grammatical gender. In a study of L1 Spanish speakers, Caffarra et al. (2014) used behavioral (both a gender agreement judgement task and a recognition memory task) and electrophysiological (EEG) methods to test the differences between transparent and opaque nouns. The findings from this study also support the idea of multiple routes to gender, which were found in previously mentioned studies espousing different theoretical approaches and methods. In their explanation, they compare the results of their multiple behavioral tests to their own ERP findings as well as previous behavioral work on multiple routes to gender (e.g., Bates et al., 1995, 1996; Gollan & Frost, 2001; Hernandez et al., 2004; Taft & Meunier, 1998).

Finally, we can also move across boundaries of linguistic modality to provide a more complete picture of grammatical gender processing. In speech, the intricacies of language processing are further complicated by conscious and unconscious mechanisms of cognitive planning as well as coordination with both articulation and paralinguistic behaviors, such as gesturing. This can often lead to differences in grammatical gender performance that contrast with established abilities of assigning and processing grammatical gender. In other words, production might not reflect comprehension, and it is important to understand why. In order to produce correct noun phrases in a language with grammatical gender, that information needs to be accessed in real time while producing the desired utterance (assuming it is not an unanalyzed chunked expression). In a review of the available research on grammatical gender in speech production (which the authors note is significantly smaller than that on grammatical gender comprehension and assignment), Wang and Schiller (2019) cover behavioral, electrophysiological, and fMRI studies that have investigated grammatical gender assignment during speech production. Based on their review, the authors conclude that, “It is generally agreed that grammatical gender is represented as a separate lexico-syntactic feature in the mental lexicon,” (Wang & Schiller, 2019, p. 5). However, despite this agreement, the authors note a number of areas of uncertainty, an important one being whether grammatical gender needs to be, or even is, accessed if grammatical gender information would have no impact on the final articulated form of the noun-phrase. That is, if the selection of one grammatical gender or another would not change the produced determiner, adjective, or noun, then it is not certain that grammatical information is actually accessed at all. This has importance regarding how and where grammatical information is linked to the lexicon. The authors continue: “Nevertheless, emerging evidence has shown

distinctive mechanisms underlying the selection of grammatical gender in Romance languages like Italian and Spanish, and Germanic languages like German and Dutch” (Wang & Schiller, 2019, p. 5). They found that this was confirmed in fMRI studies that

provide evidence for distinctive neural networks for the processing of grammatical gender and suggest that participants tend to adopt a more form-related route to access gender information in Romance languages where the gender-to-ending regularity modulates the gender effect. By contrast, participants tend to adopt a more lexically based route to access grammatical gender in Dutch and German where the noun’s morpho-phonological form is generally not strongly marked by gender. (Wang & Schiller, 2019, p. 5)

From a speech production standpoint, the findings seem to indicate that the strength of morpho/phonological cues can change the way grammatical gender is accessed. In other words, whether the anticipation of a word ending alone or the entire word is required to trigger gender assignment during oral production.

In a curious case on L1 Italian, Vigliocco et al. (1997) compared the effects of grammatical gender to “tip of the tongue (TOT) states,” (p. 314). According to the authors, “The TOT state reflects the failure to recall a word for which one has well-established knowledge” (Vigliocco et al., 1997, p. 314). Relevant to grammatical gender, a pertinent question is whether, in such a state, speakers still have access to the grammatical gender of the word or if it is blocked until the lexical entry itself is retrievable. The authors report that “speakers in a positive TOT state do have access to syntactic features of words for which they cannot yet generate a pronunciation code,” (Vigliocco et al., 1997, p. 316). This provides strong evidence for multiple routes to gender and the idea that grammatical gender information is, as stated by Wang and Schiller (2019), stored as a separate lexico-syntactic feature in the mental lexicon that is, by itself, retrievable. This finding regarding TOT states, however, returns us to the question of route to gender. I will leave this section with a question, rather than attempt any conclusion, as more research is necessary: If speakers in positive TOT states can still access grammatical gender without a lexical or morpho-phonological route, what route are they taking, and can a connectionist network explain this effect?

In this section, we have seen a number of ways that researchers think about grammatical gender, as well as ways in which they study this phenomenon, and in doing so, we crossed a number of theoretical and methodological boundaries. By moving between formal and psycholinguistic approaches to gender, we gained a deeper understanding of the structure of grammatical gender through its typological differences, its linguistic forms, and its mental representations. By contrasting findings from behavioral experiments and brain-imaging technology, we better understand how grammatical gender comprehension and production is represented by biological mental states. And by crossing between comprehension-based and production-based methods, we better understand not only how grammatical gender is processed and produced, but also the connections between processing and production generally and their broad implications for other linguistic features and human communication as a whole.

4 Grammatical Gender in SLA

Across the border between L1 and L2 research, I find both similar and unique questions about grammatical gender, and L1-L2 differences in grammatical gender provide a rich point of entry. As a base question, we would want to know how different languages themselves influence L2 grammatical gender acquisition. Linguistic systems can differ considerably in the variability, reliability, and frequency of gender cues. These differences can show the complexity that grammatical gender espouses when it is combined with other grammatical features like case, number, and gender. This can alter the relative ease of acquisition of grammatical gender. In two studies, Kempe and MacWhinney (1996, 1998) trained two groups of L1 English speakers either German or Russian as an L2 to see which of the grammatical gender systems would be more difficult to learn. German has three genders, one plural form, and four cases which combine to form 16 unique gender, number, and case combinations. Russian also has three genders, plural, and six cases, which makes 24 unique gender, number, and case combinations. One might expect German to be easier for the learners, since there are eight fewer gender, number, and case combinations to learn. However, the researchers showed that it is not simply a function of the number of combinations needed to learn, but rather an issue of cue reliability. Each gender, number, and case combination in Russian has a unique form, whereas in German, learners run into homophony/homography. The word *der* [the] illustrates this point nicely, in that it could refer to a masculine, singular noun in nominative case; a feminine, singular noun in dative or genitive case; or a plural noun in genitive case. So, when it comes to second language acquisition, the make-up of the grammatical gender system to be learned can play a significant role in ease of acquisition.

Beyond the structure of language to be learned, the type of knowledge that a person can bring from their first language(s), or transfer, also affects L2 acquisition of grammatical gender. Formal linguists, especially those interested in UG, have tested multiple theories to answer the question of transfer, including No Transfer/No Access (Epstein et al., 1996), Partial Transfer/Full Access (e.g., Minimal Trees Hypothesis, Vainikka & Young-Scholten, 1996; Failed Functional Features Hypothesis, Hawkins & Chan, 1997), and Full Transfer/Full Access (Schwartz & Sprouse, 1996). The comparison of these models had important repercussions for formal SLA, and a number of researchers attempted to find the hypothesis that best modelled learner data. In a study of L1 English learners of Arabic, Aljadani (2019) tested the Failed Functional Features Hypothesis against Full Transfer/Full Access using a grammaticality judgment task (GJT) of demonstrative pronouns. Participants were grouped by proficiency level and compared to a native speaker group. The results of the study showed that the more proficient learners did not vary significantly from the native-speaker group, whereas the less proficient group did. Based on their findings, the authors argued that since proficiency had an effect, it “could be in some way evident to the FT/FA hypothesis,” (Aljadani, 2019, p. 84). In another study using a GJT as well as a production task, Ayoun (2007) examined learnability

of grammatical gender among L1 English learners of French. Again, the authors were interested in comparing the expected results from the Failed Functional Features Hypothesis to the Full Transfer/Full Access model. Like Aljadani's (2019) study, Ayoun (2007) noted that "participants were more accurate as their level of proficiency increased," (p. 160). While both of these studies seem to provide evidence for transferability from and access to UG, from a formal perspective, it is still a messy process. This was the case in a study by Sabourin et al. (2006), which investigated learners of Dutch from multiple L1 backgrounds (German, Romance language, or English). While the results clearly bore out expected advantages (German providing a greater advantage due to its similarity with Dutch, then Romance languages because they have grammatical gender, and finally English, where grammatical gender is all but absent), the results varied by task. The authors noted that "the German group show[ed] less advantage of surface transfer for gender agreement than for gender assignment," (Sabourin et al., 2006, p. 26), which they interpret to mean that there are different effects for knowledge transfer than language use. As the authors state, "It seems that certain aspects of grammar (e.g., gender) can be learned to a high degree of accuracy but that using this knowledge remains a problem for L2 learners," (Sabourin et al., 2006, p. 26). And while proficiency played a role on some items,

an effect of L1 remains for the middle frequency items. This shows that agreement is still more difficult than assignment. Further, the English group that performs well on the assignment task only performs at chance on the agreement task. This suggests that at least for a group with no gender in their L1, gender agreement is very difficult and may be impossible to acquire. (Sabourin et al., 2006, p. 26)

The various levels of L1 overlap with the L2 grammatical system, while still influenced by proficiency and language ability, provide at least a gray-scale picture of what transfer can look like between languages, and that sharing a similar functional feature between languages can facilitate both assignment of the L2 feature, as well as, if not as much as, production of agreement patterns.

Other theoretical perspectives informed by cognitive science and developmental psychology in SLA have also taken an interest in issues of transfer and gender assignment. In a study by Carroll (1999) on beginning L2 learners of French, the researcher questioned how grammatical gender is applied at the onset of L2 acquisition. She found that learners rely on prior semantic and abstract knowledge, not just from the "objective patterns in the speech signal" (p. 38). This means that it is not just a matter of cue distribution and reliability within the input, which is a very strong driver in L1 grammatical gender acquisition. For L2 learners, the semantic relations already stored in the L1 lexicon can affect the way learners expect grammatical gender to work in an L2. She also argues that this provides "support for theories of linguistic cognition involving mediating structural representations, as well as learning theories in which conceptual information can guide grammatical development" (Carroll, 1999, p. 38). This would include Vygotskian sociocultural theory, which places an emphasis on language as a mediating tool.

SLA researchers taking cues from L1 research know that L1 readers make use of grammatical gender, but it was unclear whether L2 learners were capable of attaining L1-like processing. In order to test whether L2 learners do or are able to behave similarly, Dussias et al. (2013) compared Spanish L1 and L2 speakers' anticipatory eye movement during a reading task. They found that "late English Spanish learners revealed sensitivity to gender marking on Spanish articles similar to that found in native speakers, but this sensitivity was affected by the level of proficiency" (Dussias et al., 2013, p. 377). Thus, from a psycholinguistic perspective, we see new reading behaviors emerge through increased noticing and processing of grammatical gender that mirrors, or at least comes close to mirroring, L1 behaviors that are used in grammatical gender processing. In another study of L1 and L2 Spanish, Grüter et al. (2012) looked at advanced L2 Spanish learners with persistent production problems of grammatical gender. The authors asked whether these production problems were the result of production-specific performance or gender information retrieval in real time. To tease out these differences, participants took part in a sentence-picture matching task, an elicited production task, and a looking-while-listening task. Results showed that in the offline language task, namely the sentence-picture matching task, advanced L2 learners performed at the same level (i.e., at ceiling level) as the L1 speakers. However, for the other tasks, which were online (real-time) tasks, the L2 group did not perform as well. The authors argued that these performance differences between L1 and L2 speakers lie at the level of representation, where a more robust L1 system allows for easier real-time processing.

This question of representation of grammatical gender in an L2 system brings us back to issues of transfer, this time from a connectionist perspective. How would a connectionist model account for differences in L1/L2 grammatical gender processing? A study by Klassen (2016) looked at how gender congruency between an L1 and L2 can affect processing. To study this, the researcher asked L1 Spanish learners of German to complete a picture naming task. In Spanish, there are only two genders (masculine and feminine), while in German, there are three (masculine, feminine, and neuter). Between these two languages, then, there are some words that overlap with the same grammatical gender (*el sombrero* [the hat] is masculine in Spanish and its German counterpart *der Hut* is as well), some that are assigned the opposite gender (*la mesa* [the table] is feminine in Spanish but its German counterpart *der Tisch* is masculine), and some that are neuter in German and either masculine or feminine in Spanish (*la casa* [the house] is feminine in Spanish but its German counterpart *das Haus* is neuter). The outcomes of the picture naming task showed "faster responses in (...) L1–L2 gender-congruent nouns than for gender-incongruent ones" (Klassen, 2016, p. 24). From a connectionist perspective, this would indicate some connectivity at the lexical level based on the shared gender information which would be speeding up access. The author posits that this "show[s] that genders common to both languages are represented as L1–L2 shared gender nodes, much like what has been shown for bilinguals whose languages have symmetric gender systems" (Klassen, 2016, p. 24). In addition to this finding, the contrast between a three- and two-gender system is of note. This is because the information stored on neuter lexical entries is not "opposite" in the sense that it is

the other gender option in the L1; it is simply different information. Surprisingly, the results show that,

neuter nouns patterned similarly to L1–L2 gender-congruent nouns, illustrating that L1–L2 gender-incongruent nouns (masculine–feminine mismatches) are subject to significantly higher levels of interference in the production of bare nouns and DPs than both L1–L2 gender congruent and L2 neuter nouns. This finding suggests that nouns of different genders in the L1 and the L2 are not all subject to the same levels of interference: gender values present only in the L2 have a distinct representation that is significantly less affected by the activation of a different L1–L2 shared gender node. (Klassen, 2016, p. 24)

In sum, the level of interference from non-congruence is highly correlated to L1/L2 similarities and differences, in addition to issues of proficiency, as was investigated in other studies on processing described earlier.

The various borders crossed in SLA reflect many of the borders crossed broadly in linguistics as discussed in the previous section. Of particular note here, though, is that the relative size of the fields does not need to be similar for the benefits of boundary crossing to be prevalent. In some ways, SLA gains an immense amount of information by crossing into theories, approaches, and findings from formal linguistics, developmental psychology, L1A, and the like, but the reverse is also true. These other fields also gain an immense amount of insight by crossing into the boundaries of more niche fields to see how their general theories play out in particular contexts and either provide evidence to strengthen their claims or force them to reconcile their theories with new evidence from specific contexts.

5 L2 Grammatical Gender Instruction

Based on the previous review, it is clear that grammatical gender is complex and often difficult to learn, although this is relative based on the L2 grammatical gender system and any opportunities for transfer from previously learned L1/L2s. Many studies provide strong evidence for difficulties learning an L2 grammatical gender system. In one study on L2 German, Walter and MacWhinney (2015) surveyed students majoring in German at US universities in their final year of their studies. They found that even among students who were about to receive a bachelor's degree in German, control over grammatical gender assignment varied considerably, with many participants only knowing approximately 50 percent of the genders of the nouns they were asked about, and there was a fairly clear lack of understanding of cues that could have provided them with information about the correct gender, even if they did not recognize the word. Knowing that grammatical gender is difficult to learn means that it is important to find evidence-based methods that can support learning. Based on my experience and interactions with fellow teachers and researchers, most pedagogues emphasize explicit instruction, awareness raising, and other fairly straight-forward approaches based on research and practices with other grammatical features. But are there differences that come about from these proposed methodologies? Is there an effect of instruction on acquisition? Instructed

SLA, as an area focused primarily on this question, would want to know how the boundary between instructed and non-instructed learners differs in their attainment, understanding, and use of grammatical gender. And within the classroom, the boundaries between different teaching methods might also play a part in this development. For example, how do comprehension-based activities differ from and production-based activities in the way they change how learners acquire, process, and use grammatical gender? Keppen et al. (2021) asked this very question and found that, for beginning L2 learners at least, production-based activities were superior for helping students on comprehension and production tasks.

In a study that utilized a cognitive tutor based on psycholinguistic principles, Presson et al. (2014) provided learners with individualized feedback on hundreds of French nouns. In each trial, learners were informed whether they were correct or not, and if incorrect, they were provided with individualized feedback about their errors. In some cases, learners received explicit rules about gender assignment in French in addition to corrective feedback, while other learners only received corrective feedback or feature focusing. The group that received explicit instruction outperformed the other two groups, although they found no effect for more versus less frequent words.

In addition to developing technological solutions to overcome the deficit in L2 gender acquisition, some researchers have tried to utilize the (fairly weak, but existent) link between sex and gender in instruction. Most of these methods have been focused on providing semantically associated information, such as objects that are stereotypically or culturally more associated with a particular sex, via mnemonic devices. Color-coding, especially, has been the focal point of textual enhancement approaches because there is at least some link between many gender systems and sex, so traditional cultural connections between some colors (e.g., pink for feminine and blue for masculine in a Western context) seemed like possible avenues to highlight bare textual information for learners and provide some additional support for gender awareness during reading. This theory was tested in studies by Desrochers (e.g., Desrochers, 1982; Desrochers et al., 1989) in the mid-1980s and early 1990s, and has received some renewed attention in recent years, especially from a deeper understanding of connectionist frameworks, where instead of just one (color) link between the gender and sex, additional semantic connections are also made. These include, for example, the sex of the actors who use or are pictured with similarly gendered nouns, as well as voiced productions of words with overlapping gender/sex categorization, i.e., a male voice actor for masculine nouns and a female voice actor for feminine nouns. Two recent studies, one by Dias de Oliveira Santos (2015) and one by Arzt and Kost (2016), both tested the effects that different combinations of these mnemonic devices can have on gender acquisition. Each study compared a control group to different conditions of color/voice/image pairings. In both cases, the researchers did not find a significant immediate effect for this type of instruction, but Arzt and Kost (2016) did find a delayed effect for both forms of visual enhancement on retention of gender information.

5.1 Grammatical Gender as Part of a Complex Grammatical-Functional System

Based on all of the issues presented with instructing grammatical gender, it is clear that simply providing learners with which genders map to which nouns is insufficient. Even many of the teaching strategies that try to bring the form to the forefront, such as visual enhancements, mnemonics, explicit instruction, and input processing approaches, seem to have little effect on the long-term acquisition of grammatical gender. I argue that one of the major reasons for the lack of both uptake and understanding of grammatical gender is the relative dearth of approaches that link this grammatical feature to its functionality, and therefore relevance, within a complex grammatical system. The function that grammatical gender plays is variable based on language, but let's take the case of German grammatical gender as one example. As previously mentioned, German grammatical gender overlaps with case and number information and is often homophonic and homographic. This leads to extensive problems beyond simply mapping gender to nouns, as the cues needed to do so are unreliable. So, rather than understanding grammatical gender as something that needs to be learned for correctness, approaches such as sociocultural theory and systemic functional linguistics put the function of the feature at the forefront rather than the form. Each language's grammatical gender feature might play a different functional role, but in German, it is intertwined with the case marking system and therefore essential knowledge to unravel sentential role. Because role is marked by case and not syntactic order, word order is very flexible in German. Thus, the functions of movement and topicalization can be used to motivate control of noun phrase declension in German. For example, if you ask a German speaker who they saw yesterday, they can easily respond "Meinen Freund sah ich." [my friend saw I], just as easily as they can respond "Ich sah meinen Freund." [I saw my friend], with the first version even being preferable because the placement of the object in the first position, marked as such by the *-en* on the possessive adjective *mein*, emphasizes and highlights the answer to the question that was actually posed.

In Walter and van Compernelle (2017) the authors enact this through concept-based instruction (CBI) (Negueruela, 2003). CBI, as defined in van Compernelle and Henery (2014),

is an approach to L2 pedagogy that centers on promoting the internalization of categories of meaning as psychological tools that mediate L2 communication. CBI draws on Vygotsky's (1986) analysis of the development of scientific thinking in formal schooling and expands and adapts Gal'perin's (1989, 1992) systemic-theoretical instruction and Davydov's (2004) germ-cell model approach to teaching subjects such as mathematics (see, e.g., Stetsenko & Arievidtch, 2010). In L2 CBI, categories of meaning are presented as abstract, systematic concepts that "are semantically driven, recontextualizable, and agentive." (p. 72)

In Walter and van Compernelle (2017), the authors, working with the framework of CBI (also concept-based language instruction, C-BLI), taught grammatical gender and case through the concepts of movement and topicalization via animated slides

and in-class enrichment activities where students practiced creating their own topicalized sentences. After comparing pre- and post-tests on a subject-identification task with reflective questions, “the CBI enrichment program not only improved learners’ scores, but also helped them to develop conscious, declarative knowledge of how the language works” (Walter & van Compernelle, 2017, p. 81). This is in contrast with previous mnemonic studies which did not show gains on immediate posttests.

In an extension of this work, Walter (2020) united this teaching framework with a cognitive-tutor implementation similar to the one described in Presson et al. (2014). In this study, four high school second-year German classes were divided into two groups. Two classes received a CBI approach with post-lesson practice on hundreds of sentence-level and noun-phrase-level exemplars, and the other two classes received traditional, text-book style explicit information before beginning with the same exemplar training. The researcher found learning in both groups, but the types of knowledge gained diverged. The more traditional approach with the large amount of corrective feedback and practice resulted in small to moderate gains in both passive and productive tasks. The CBI program, on the other hand, had larger effects for a smaller number of categories, specifically the inclusion and accuracy of attributive adjective endings and the picture-sentence matching task. Also, reflective comments provided in the post-test, which asked specifics about the grammatical gender system of German, showed differences in the type of vocabulary used by learners, with the CBI group focusing much more on the function of grammar over linguistic terminology.

The different approaches in Instructed SLA derived from multiple SLA theories, as well as other pedagogical traditions, has led to an array of methods designed to best instruct what we know about grammatical gender, how it is learned and acquired, and how best to teach it. Currently, there are a number of competing methods. Interpreting each of the learning outcomes from a single perspective is unlikely to provide a true representation of the different effects of these methods, and therefore boundary crossing will likely be needed to understand no whether one approach is “better” than another, but how each method affects learning differently. It is highly likely that a new method or combination of methods works best. One that only becomes visible when one takes an approach that is informed by diligent border crossing and interdisciplinary cooperation.

6 Boundaries Crossed and Those Still to Cross

From this review of research on grammatical gender, it is clear that fully understanding this phenomenon requires a multidisciplinary approach that crosses a number of boundaries. It is important to cross language boundaries to understand how grammatical gender is instantiated in one language versus another, and how the differences in those systems affect the way grammatical gender is assigned and used. It is important to cross semantic boundaries between terms like grammatical gender and sex, and grammatical gender and noun class, and see how the use of

certain terminology can affect the way people perceive these linguistic structures. It is important to cross methodological boundaries because different methods and theories can provide us a more complete picture of grammatical gender. It is important to cross teleological boundaries to understand how language change can affect the structure and use of grammatical gender over phylogenetic, ontogenetic, and microgenetic timescales, as well as through historical and cultural processes. And finally, it is important to cross pedagogical boundaries in an attempt to make this difficult-to-acquire linguistic feature as learnable as possible.

In the spirit of this book and the Tuckerian impact of boundary crossing, it is essential that researchers from different theoretical and methodological backgrounds engage with those from other perspectives. I believe that by providing spaces for this to occur, scholars would better understand the reasons behind others' perspectives. It is also imperative that researchers allow themselves to become boundary crossers. This means that if another perspective provides a better model to support the data, that person must be willing to either accept that other model as more accurate or be willing to alter their model to become the more accurate model of the data. I can see this from Dr. Tucker's initial assessment of French linguistics described at the beginning of this chapter. Boundary crossing is a humbling experience. One presents oneself, an expert over here, as somewhat of a neophyte over there. In these moments of humility, one can gain a new appreciation for the work and perspective of others and find oneself in a position to listen and learn, rather than to profess and explain.

It is also critical that we allow for more boundary crossing between the theoretical/experimental study of linguistic features and pedagogical/instructional implementation (cf. how Sects. 1 and 2 crosscut in this volume). I see a particular need with studies on grammatical gender to understand how the terminology we use impacts learners' assumptions and development through a combination of work from research on linguistic relativism, psycholinguistics, and instructed SLA. This will require more direct communication between researchers and teachers, and, again, providing a space where teachers see themselves as integral to our understanding of how language acquisition happens in the classroom, and where researchers see themselves as sources for classroom innovation.

References

- Aikhenvald, A. (2004). *Gender and noun class*. Walter de Gruyter.
- Aljadani, A. S. (2019). The acquisition of grammatical gender in Arabic demonstratives by English native speakers. *International Journal of Contemporary Education*, 2(1), 78–86.
- Arnon, I., & Ramscar, M. (2012). Granularity and the acquisition of grammatical gender: How order-of-acquisition affects what gets learned. *Cognition*, 122(3), 292–305.
- Arzt, J., & Kost, C. (2016). Effect of different teaching techniques on the acquisition of grammatical gender by beginning German second language learners. *Die Unterrichtspraxis/Teaching German*, 49(1), 1–23.

- Ayoun, D. (2007). The second language acquisition of grammatical gender and agreement. *French Applied Linguistics*, 16, 130–170.
- Bates, E., Devescovi, A., Pizzamiglio, L., D'amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, 57(6), 847–862.
- Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, 58(7), 992–1004.
- Bender, A., Beller, S., & Klauer, K. C. (2011). Grammatical gender in German: A case for linguistic relativity? *Quarterly Journal of Experimental Psychology*, 64(9), 1821–1835.
- Blom, E., Polisenska, D., & Weerman, F. (2008). Articles, adjectives and age of onset: The acquisition of Dutch grammatical gender. *Second Language Research*, 24(3), 297–331.
- Boroditsky, L. (2001). Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology*, 43(1), 1–22.
- Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 61–79). The MIT Press.
- Caffarra, S., Janssen, N., & Barber, H. A. (2014). Two sides of gender: ERP evidence for the presence of two routes during gender agreement processing. *Neuropsychologia*, 63, 124–134.
- Caffarra, S., Siyanova-Chanturia, A., Pesciarelli, F., Vespignani, F., & Cacciari, C. (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, 52(8), 1019–1030.
- Carroll, S. E. (1999). Input and SLA: Adults' sensitivity to different sorts of cues to French gender. *Language Learning*, 49(1), 37–92.
- Carstens, V. (2010). Implications of grammatical gender for the theory of uninterpretable features. In M. T. Putnam (Ed.), *Exploring crash-proof grammars* (pp. 31–58). John Benjamins.
- Davydov, V. (2004). *Problems of developmental instruction: A theoretical and experimental psychological study*. (P. Moxay, Trans.). Akademyia Press.
- De Resende, N. C. A., Mota, M. B., & Seuren, P. (2019). The processing of grammatical gender agreement in Brazilian Portuguese: ERP evidence in favor of a single route. *Journal of Psycholinguistic Research*, 48(1), 181–198.
- Desrochers, A. (1982). *Effects of an imagery mnemonic on the acquisition and retention of French article-noun pairs*. Unpublished Doctoral Dissertation. University of Western Ontario.
- Desrochers, A., & Paivio, A. (1990). Le phonème initial des noms inanimés et son effet sur l'identification du genre grammatical. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 44(1), 44–57.
- Desrochers, A., Gélinas, C., & Wieland, L. D. (1989). An application of the mnemonic keyword method to the acquisition of German nouns and their grammatical gender. *Journal of Educational Psychology*, 81(1), 25–32.
- Dias de Oliveira Santos, V. (2015). Can colors, voices, and images help learners acquire the grammatical gender of German nouns? *Language Teaching Research*, 19(4), 473–498.
- Dominguez, A., Cuetos, F., & Segui, J. (1999). The processing of grammatical gender and number in Spanish. *Journal of Psycholinguistic Research*, 28(5), 485–498.
- Dussias, P. E., Kroff, J. R. V., Tamargo, R. E. G., & Gerfen, C. (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, 35(2), 353–387.
- Epstein, S. D., Flynn, S., & Martohardjono, G. (1996). Universal grammar and second language acquisition: The null hypothesis. *Behavioral and Brain Sciences*, 19(4), 746–758.
- Gal'perin, P. I. (1989). Organization of mental activity and the effectiveness of learning. *Soviet Psychology*, 27(3), 65–82.
- Gal'perin, P. I. (1992). Stage-by-stage formation as a method of psychological investigation. *Journal of Russian & East European Psychology*, 30(4), 60–80.
- Gollan, T. H., & Frost, R. (2001). Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, 30(6), 627–651.

- Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191–215.
- Hawkins, R., & Chan, C. Y.-H. (1997). The partial availability of universal grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research*, 13(3), 187–226.
- Hernandez, A. E., Kotz, S. A., Hofmann, J., Valentin, V. V., Dapretto, M., & Bookheimer, S. Y. (2004). The neural correlates of grammatical gender decisions in Spanish. *Neuroreport*, 15(5), 863–866.
- Hohlfeld, A. (2006). Accessing grammatical gender in German: The impact of gender-marking regularities. *Applied PsychoLinguistics*, 27(2), 127–142.
- Kay, P., & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist*, 86(1), 65–79.
- Kempe, V., & MacWhinney, B. (1996). The crosslinguistic assessment of foreign language vocabulary learning. *Applied PsychoLinguistics*, 17(2), 149–183.
- Kempe, V., & MacWhinney, B. (1998). The acquisition of case marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, 20, 543–587.
- Keppenne, V., Hopman, E. W., & Jackson, C. N. (2021). Production-based training benefits the comprehension and production of grammatical gender in L2 German. *Applied PsychoLinguistics*, 42(4), 907–936.
- Klassen, R. (2016). The representation of asymmetric grammatical gender systems in the bilingual mental lexicon. *Probus*, 28(1), 9–28.
- Köpcke, K.-M., & Zubin, D. A. (1984). Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klassifikation. *Linguistische Berichte*, 93, 26–50.
- Köpcke, K.-M., & Zubin, D. A. (1996). Prinzipien für die Genuszuweisung im Deutschen. In E. Lang & G. Zifonun (Hrsg.), *Deutsch typologisch. Jahrbuch des Instituts für Deutsche Sprache 1995* (pp. 473–491). Walter de Gruyter.
- Kousta, S.-T., Vinson, D. P., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 843–858.
- Kramer, R. (2014). Gender in Amharic: A morphosyntactic approach to natural and grammatical gender. *Language Sciences*, 43, 102–115.
- Matthews, C. (2010). On the nature of phonological cues in the acquisition of French gender categories: Evidence from instance-based learning models [DP- internal information structure]. *Lingua*, 120(4), 879–900.
- Mills, A. E. (2012). *The acquisition of gender: A study of English and German*. Springer-Verlag.
- Negueruela, E. (2003). *Systemic-theoretical instruction and l2 development: A sociocultural approach to teaching-learning and researching l2 learning*. Unpublished Doctoral Dissertation. The Pennsylvania State University.
- Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th annual meeting of the cognitive science society* (pp. 928–933). Cognitive Science Society.
- Presson, N., MacWhinney, B., & Tokowicz, N. (2014). Learning grammatical gender: The use of rules by novice learners. *Applied PsychoLinguistics*, 35(4), 709–737.
- Sabourin, L., Stowe, L. A., & De Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22(1), 1–29.
- Samuel, S., Cole, G., & Eacott, M. J. (2019). Grammatical gender and linguistic relativity: A systematic review. *Psychonomic Bulletin & Review*, 26(6), 1767–1786.
- Schwartz, B. D., & Sprouse, R. A. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, 12(1), 40–72.
- Stetsenko, A., & Arievidt, I. M. (2010). Cultural-historical activity theory: Foundational worldview and major principles. In J. Martin & S. Kirschner (Eds.), *The sociocultural turn in psychology: The contextual emergence of mind and self* (pp. 231–253). Columbia University Press.

- Taft, M., & Meunier, F. (1998). Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research*, 27(1), 23–45.
- Tucker, G. R. (1967). *The French speaker's skill with grammatical gender: An example of rule-governed behavior*. [Unpublished doctoral dissertation]. McGill University.
- Tucker, G. R., Lambert, W. E., Rigault, A., & Segalowitz, N. (1968). A psychological investigation of French speakers' skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, 7(2), 312–316.
- Vainikka, A., & Young-Scholten, M. (1996). Gradual development of L2 phrase structure. *Second Language Research*, 12(1), 7–39.
- van Compernelle, R. A., & Henery, A. (2014). Instructed concept appropriation and L2 pragmatic development in the classroom. *Language Learning*, 64(3), 549–578.
- Vigliocco, G., Antonini, T., & Garrett, M. F. (1997). Grammatical gender is on the tip of Italian tongues. *Psychological Science*, 8(4), 314–317.
- Vygotsky, L. (1986). *Thought and language*. The MIT Press.
- Walter, D. R. (2020). The acquisition of German declension in additive and concept-based approaches to instruction via computer-based cognitive tutors. *Language, Interaction and Acquisition*, 11(1), 130–162.
- Walter, D. R., & MacWhinney, B. (2015). US German majors' knowledge of grammatical gender. *Die Unterrichtspraxis/Teaching German*, 48(1), 25–40.
- Walter, D. R., & van Compernelle, R. A. (2017). Teaching German declension as meaning: A concept-based approach. *Innovation in Language Learning and Teaching*, 11(1), 68–85.
- Wang, M., & Schiller, N. O. (2019). A review on grammatical gender agreement in speech production. *Frontiers in Psychology*, 9, 27–54.
- Zubin, D. A., & Köpcke, K.-M. (1981). Gender: A less than arbitrary grammatical category. *Chicago Linguistic Society*, 17, 439–449.
- Zubin, D. A., & Köpcke, K.-M. (1986). Gender and folk taxonomy: The indexical relation between grammatical and lexical categorization. In C. Craig (Ed.), *Noun classes and categorization* (pp. 139–180). John Benjamins.

Non-Expert Raters' Scoring Behavior and Cognition in Assessing Pragmatic Production in L2 Chinese



Shuai Li, Xian Li, Yali Feng, and Ting Wen

Abstract This chapter reports on a study investigating non-expert raters' scoring behavior and cognitive processes involved in evaluating speech acts and pragmatic routines in L2 Chinese. Pragmatic production data were collected from 51 American learners of Chinese, who completed a 12-item oral Discourse Completion Test (DCT). The learners were divided into 15 groups, each including the same six learners and three different learners. A total of 101 non-expert, native Chinese raters evaluated the oral productions of one learner group and were encouraged to verbalize their scoring rationale. Results showed that, although the raters varied significantly in scoring severity, their scoring behaviors were consistent, with very limited instances of scoring bias. Qualitative analysis based on 2753 verbal protocols revealed that the raters predominantly oriented towards criteria related to holistic meaning expression in assessing speech acts and routines. They prioritized criteria related to linguistic expressions (notably those concerning vocabulary knowledge) in evaluating pragmatic routines, and they paid more attention to criteria related to interactional skills in assessing speech acts. Boundary crossing implications are discussed in relation to pragmatics assessment and L2 Chinese teaching.

Keywords Pragmatics assessment · L2 Chinese · Rater cognition · Scoring behavior

S. Li (✉)

Department of World Languages and Cultures, Georgia State University, Atlanta, GA, USA
e-mail: sli12@gsu.edu

X. Li · Y. Feng

Department of Applied Linguistics/ESL, Georgia State University, Atlanta, GA, USA
e-mail: xli66@gsu.edu; yfeng9@gsu.edu

T. Wen

College of Chinese Studies, Beijing Language and Culture University, Beijing, China
e-mail: wenting@blcu.edu.cn

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_4

1 Introduction

With the publication of the seminal work by Hudson et al. (1995), second language (L2) pragmatics assessment evolved to cross the boundaries of research on interlanguage pragmatics and on L2 performance assessment. Informed by pragmatics theories and research on interlanguage pragmatics, the field has mainly focused on developing instruments for assessing pragmatic competence, leading to expanded construct coverage in pragmatics assessment (Taguchi & Roever, 2017). More recently, thanks to the influence from L2 performance assessment, the field's research agenda has pluralized. Researchers have increasingly paid attention to contingent factors that may affect assessment outcomes, including rater behavior and cognition (e.g., Liu & Xie, 2014; Youn, 2007), rating scale functioning (e.g., Chen & Liu, 2016; Li et al., 2019), and differential item functioning (e.g., Roever, 2007), to name just a few. Among such factors, rater behavior and cognition have attracted much research attention, which reflects sustained interest in this topic in the larger field of L2 performance assessment (for a review, see Han, 2016). Meanwhile, methodological boundary crossing has also characterized the field's development. Whereas early-stage studies adopted a strong psychometric paradigm, more recent research has incorporated additional methodological paradigms such as Conversation Analysis and discursive pragmatics (e.g., Youn, 2015; Walters, 2007).

In L2 pragmatics assessment, researchers typically develop *a priori* rating criteria based on theorizations of pragmatic competence, and recruit what we refer to as *expert raters* who are trained in a closely related academic field (e.g., pragmatics, applied linguistics). Such expert raters have been found to exhibit considerable variability in scoring severity (e.g., Liu & Xie, 2014; Youn, 2007); in interpretation of the substantive meaning of rating criteria (e.g., Li et al., 2019); in scoring bias towards examinees, pragmatic features, and/or assessment items (e.g., Youn, 2007); and in prioritization of certain rating criteria over others (e.g., Taguchi, 2011; Walters, 2007). However, the common practice in the field of only having expert raters evaluate pragmatic performance can be problematic because other potentially relevant stakeholders are left out. Such stakeholders include, for example, native speakers of the target language who are not equipped with the kind of academic training and/or teaching experiences that expert raters have. Such native speakers are the people that L2 learners are supposed to interact with outside the language classroom (e.g., consider the study abroad context), and they are the people who evaluate learners' performance in real world contexts. Hence, it is critical to include such *non-expert raters* into pragmatics assessment.

To date, with only the exception of Taguchi (2011), very little is known about how non-expert raters would assess L2 pragmatic performance and what evaluation criteria they would adopt. Answers to these questions would have boundary-crossing implications: they would allow us to gauge the generalizability of existing findings regarding expert raters' scoring behavior and cognition; such information would also inform L2 instruction and learning by understanding which aspects of linguistic performance are deemed important by potential stakeholders. This study intends

to contribute to this line of research by investigating how non-expert native Chinese speakers evaluate the production of speech acts and pragmatic routines in L2 Chinese.

2 Literature Review

In this section, we start with a brief review of the rating criteria used to evaluate L2 pragmatic performance. We then discuss quantitative studies on raters' scoring behavior. In the spirit of paradigmatic boundary crossing, the review of quantitative research is complemented by a discussion of qualitative studies on rater cognition, because raters' cognitive processes during scoring have been found to be related to their scoring behavior. This section ends with a critique of the existing literature from a boundary crossing perspective.

2.1 Rating Criteria in L2 Pragmatics Assessment

Rating scales with descriptors of evaluation criteria have been widely used to assess L2 pragmatic performance (Taguchi & Li, 2021). Such criteria have developed over time to reflect the evolving theorizations of pragmatic competence (for a recent review, see Li, 2021). In the early stage of rating criteria development, researchers resorted to the understanding of pragmatic competence as consisting of pragmalinguistic and sociopragmatic components. The former refers to the connections between linguistic forms and their pragmatic functions, and the latter concerns the sociocultural rules underlying linguistic behavior of a particular speech community (Leech, 1983; Thomas, 1983). Pragmatics rating criteria during that period mainly addressed considerations of appropriateness, directness, and politeness; other criteria adopted in the early stage of pragmatics assessment research included realization of communicative intention, use of formulaic expressions (for assessing pragmatic routines), and amount of speech/information (e.g., Hudson et al., 1995; Liu, 2006).

While the broadly defined notion of appropriateness has remained in all pragmatics assessment research since Hudson et al.'s (1995) foundational project, researchers have later incorporated additional criteria into assessment. One notable addition was the inclusion of linguistic accuracy in evaluating pragmatic performance (e.g., Chen & Liu, 2016; Grabowski, 2013; Li et al., 2019; Taguchi, 2012), which reflects the close relationship between grammatical and pragmatic competencies (Bardovi-Harlig, 2003). More recent, and significant, additions to pragmatics assessment criteria have been informed by theorizations of interactional competence (e.g., Young, 2011) and discursive pragmatics (Kasper, 2006). Under these perspectives, pragmatic competence is not considered as an individual trait, but rather as an ability that emerges in the process of co-constructing meaning in interaction (Taguchi, 2019). Interactional skills such as turn-taking, topic management,

and repair are thus critical for understanding pragmatic competence and have been incorporated into pragmatics assessment (Timpe, 2013; Youn, 2015).

As discussed above, appropriateness, linguistic accuracy, and interactional skills constitute the major dimensions of pragmatics assessment criteria. Because speech acts have been the main focus in the field, existing rating criteria are best at serving the purpose of assessing speech acts rather than other pragmatic constructs (e.g., pragmatic routines). In actual assessment practice, raters are typically provided with a set of *a priori* rating criteria. While such rating criteria can orient raters to the major dimensions to consider during the scoring process, they do not delineate specific linguistic features that may lead to a higher or lower score. An example is Li et al.'s (2019) study that assessed the production of compliment responses, refusals, and requests. The score descriptor of Band 4 (there were six scoring levels) reads “target communicative function somewhat realized; expression somewhat appropriate for a given scenario (e.g., verbosity, somewhat more direct and/or indirect than needed, use of uncommon semantic formula) as judged by native speaker raters; syntactic and/or lexical errors tend to interfere with meaning and/or appropriateness” (p. 293). Such general descriptions of benchmark performance for this score band leaves plenty of room for interpretation by raters. For example, exactly what linguistic features in examinees’ productions constitutes “somewhat appropriate” performance may be quite different across raters’ minds. Variability in rater cognition may, in turn, influence their scoring behavior (discussed below).

2.2 *Raters’ Scoring Behavior in Assessing L2 Pragmatics: Quantitative Studies*

Quantitative research on raters’ scoring behavior has mainly focused on understanding whether raters exhibit similar or different levels of severity, whether they perform scoring consistently (e.g., being consistent in scoring severity), and whether they demonstrate any bias in scoring (i.e., being particularly harsh or lenient for certain examinees, assessment items, and/or pragmatic features). These issues are typically investigated by using the Rasch model, which is a psychometric model widely adopted in L2 performance assessment (McNamara et al., 2019). Based on raw scores, the Rasch model estimates rater severity, examinee ability, and difficulty of test items on a *logit scale*. The logit scale is an interval scale centered at the zero point and extending to positive and negative infinity. The measurement unit on the logit scale is called a *logit*. A larger (or positive) logit value indicates greater severity of raters in scoring, higher ability of examinees, and a higher difficulty level of assessment items, and vice versa. Moreover, the Rasch model outputs separation indices to indicate the number of statistically distinct levels of rater severity, examinee ability, and item difficulty. It also calculates *fit* statistics (called *Mean Square*, or *MnSq*) to reveal the extent to which the response patterns of individual raters, examinees, and test items conform to the model’s expectations. An acceptable range

of MnSq to indicate good model fit is 0.5–1.5 (Wright & Linacre, 1994), although some researchers have also used the more conservative range of 0.7–1.3 (e.g., Liu, 2006). Finally, the Rasch model allows bias/interaction analysis among the variables. For example, it can tell whether a rater gives particularly harsh scores to specific examinees and/or specific items.

Existing studies have predominantly focused on scoring behavior of what we previously referred to as *expert raters*. For example, Youn (2007) studied three expert raters' behavior in scoring speech act production (including apologies, refusals, and requests) in L2 Korean based on Hudson et al.'s (1995) rating criteria (discussed in the previous section). The raters were all native Koreans with graduate training in applied linguistics, and two of them also had relevant teaching experience. Results showed that the raters' scoring behavior conformed to the expectations of the Rasch model, but they significantly differed in scoring severity. The raters showed different bias patterns in assigning scores to individual examinees. Similar findings were reported by Liu and Xie (2014), who recruited both native and non-native English speaker raters (who were all college English instructors) to evaluate written production of apologies by Chinese EFL learners. The raters showed biases in scoring certain examinees, which was likely due to differences in prioritizing certain criteria during the scoring process. For example, some raters considered grammatical knowledge to be critical, but others attached more importance to how examinees realized apologies. Raters' differential interpretation of rating criteria was also reported in Li et al.'s (2019) study, where two expert native speaker raters with shared academic, cultural, linguistic, and professional backgrounds evaluated speech act production (including compliment responses, refusals, and requests) in L2 Chinese.

Collectively, findings of the small number of existing studies suggest that expert raters, in assessing speech act production with a set of *a priori* rating criteria, are generally able to assign scores consistently (i.e., their scoring patterns meet the expectation of the Rasch model), but they often show considerable variation in scoring severity and may exhibit scoring biases towards examinees or assessment items. It is unclear whether non-expert raters without relevant academic training or instructional experience would demonstrate similar scoring behavior. Moreover, raters' scoring bias and varied severity in scoring may be related to their individualized cognitive processes, as demonstrated in Liu and Xie's (2014) study. Variability in rater cognition is an issue often examined in qualitative studies, which are reviewed in the next section.

2.3 Variability in Rater Cognition: Qualitative Studies

Research on rater cognition in L2 pragmatics assessment typically analyses raters' protocols detailing their cognitive processes during scoring in order to investigate which aspect(s) of examinee performance they attend to. The small body of literature has focused on the effects of native speaker status and varied native language

backgrounds on rater cognition, and only one study included non-expert raters (Taguchi, 2011).

Two studies compared rater cognition between native and non-native expert raters (Alemi & Tajeddin, 2013; Walters, 2007). In Walters' (2007) study, two expert raters (native and non-native English speakers) trained in conversation analysis (CA) were recruited. Both raters evaluated one ESL learner's role play of two CA-informed constructs (i.e., assessment, pre-sequence) and one speech act (i.e., compliment) in terms of the level of realization. The two raters discussed discrepancies in their ratings through a series of dialogues, which revealed considerable differences in how they interpreted the same performance. Whereas the non-native rater (who shared the same native language as the learner) cited L1 transfer as a possible explanation of the learners' non-native-like performance, the native rater relied on his intuition for evaluation. Moreover, the non-native speaker also paid attention to fluency and clarity in pronunciation, but the native speaker did not.

While it may be difficult to attribute Walters' findings to native status because of the small sample size of his study, Alemi and Tajeddin's (2013) study demonstrated rater variability between native and non-natives with a larger group of rater participants. The researchers recruited 50 native English raters (who were ESL faculty) and 50 non-native raters (who received M.A. training in applied linguistics and had multiple years of teaching experience) to evaluate refusals in L2 English elicited through a written DCT. The raters evaluated overall appropriateness of the refusal responses and wrote down their scoring rationale. Results showed that the non-native raters were more lenient than their native counterparts. Regarding rater cognition, while the native raters resorted to 11 criteria during the scoring process, the non-native raters only referred to six criteria. Moreover, the two rater groups differed in their predominant evaluation criteria: whereas politeness was the most important consideration among the non-native raters, provision of appropriate reasoning and explanation was the leading criterion among the native raters.

As the aforementioned studies show, expert raters sometimes employ criteria that may not be incorporated in theory-informed pragmatics assessment literature (see the first section of this literature review), such as fluency and pronunciation. This tendency is more clearly shown in Sydorenko, Maynard, and Guntly's (2014) study. Three expert raters (with ESL teaching experience and familiarity with the speech act literature) listened to ESL learners' oral production of multiple-turn requests, evaluated the level of overall appropriateness, and explained their scoring rationale. Results showed that the raters paid attention to the sequential organization of requests, noting the follow-up moves (e.g., thanking, closing) after a request was delivered. The raters also considered the specific contexts in which request utterances occurred, as well as intonational patterns, repetitiveness of speech, and cultural misunderstanding.

While the above studies all focused on expert raters, Taguchi's (2011) study is the only one that included non-expert raters. Similar to Sydorenko et al.'s findings, Taguchi's non-expert raters also paid attention to various aspects of speech act production during scoring, and there were considerable variations in individual raters' cognition. Taguchi's raters were all native English speakers but differed in cultural

backgrounds: there were one African American male, one Australian white male and one female, and one Japanese American female. The raters evaluated appropriateness of two speech acts (i.e., request and opinion) in L2 English, and shared their scoring rationales through individual introspective interviews. Results showed that the raters tended to focus on different dimensions of learner performance as the basis for scoring: some prioritized linguistic forms but others paid more attention to semantic content and strategies. The raters also varied in their level of tolerance for the same aspect of performance. Finally, some raters also resorted to personal experience to support scoring decisions.

2.4 A Boundary Crossing Critique of the Literature, and This Study

Our literature review so far has shown that the mainstream practice of L2 pragmatics assessment typically relies on expert raters and adopts theory-informed, *a priori*, rating criteria that mainly focus on broadly defined dimensions of appropriateness, linguistic accuracy, and interactional skills. Existing studies suggest that such pre-determined rating criteria are often open to individualized interpretations when expert raters evaluate specific instances of pragmatic performance. During the scoring process, expert raters are also likely to prioritize certain criteria over others, may attend to features that are not typically assessed in the literature (e.g., fluency), and factor in their personal experiences and/or expectations. Such variability in rater cognition may influence raters' scoring behavior (e.g., severity in scoring, bias in scoring).

From a boundary crossing perspective, several issues need to be addressed. First, the field's predominant focus on expert raters (except for Taguchi's study discussed above) artificially creates a "rater eligibility boundary" in pragmatics assessment based on professional training/knowledge, which underestimates the importance of other potential stakeholders of L2 pragmatics assessment. As Sydorenko et al. (2014) contended, criteria for assessing L2 pragmatic performance should not come exclusively from experts in pragmatics research or experienced language professionals, but also from people who are most likely to interact with the targeted examinee population. We would argue that such people include what we previously referred to as *non-expert raters*, who are not savvy in linguistics or pedagogical theories and may not have rich experiences in interacting with L2 speakers. Such non-expert raters should be included in the practice of pragmatics assessment because, arguably, they are the most likely interlocutors for L2 learners outside the classroom. Second, the field's almost exclusive focus on speech acts (except for Walters' study reviewed above) in understanding rater behavior and cognition reflects and reinforces a "target construct boundary" that is still in place in the larger field of interlanguage pragmatics, as speech acts have long been the most extensively researched pragmatic feature (Taguchi & Roever, 2017). Such a target construct boundary in L2 pragmatics assessment research tends to restrict our

understanding of whether and how raters may adjust their cognitive processes according to different pragmatic features. Incorporating pragmatic features in addition to speech acts is thus in order. Finally, existing research on rater cognition has exclusively focused on English as the target language. This “target language boundary” needs to be crossed given the considerable cultural and linguistic variations among world languages.

Inspired by the boundary crossing spirit of G. Richard Tucker (see Zhang & Miller, this volume), the present study aimed to address the aforementioned issues by focusing on the rating behavior and cognition of non-expert raters who evaluated both speech acts and pragmatic routines in L2 Chinese. We adopted quantitative and qualitative methodological approaches to answer the following research questions.

RQ1. What are the patterns of non-expert raters’ scoring behavior?

RQ2. What criteria do non-expert raters adopt to evaluate speech acts and pragmatic routines?

3 Method

3.1 Examinees

Examinee data came from 51 American learners of Chinese recruited from a study abroad program in China. There were 22 males and 29 females, with a mean age of 20.41 years ($SD = 0.96$). At the time of data collection, the examinees were just starting their study abroad semester. Prior to going abroad, they had received, on average, 2.22 years of formal instruction in Chinese ($SD = 1.18$). The examinees took the New HSK for placement. The HSK test is a standardized Chinese proficiency test suite consisting of separate tests for six proficiency levels for the written part (tapping listening, reading, and writing) as well as separate tests for three oral proficiency levels (elementary, intermediate, and advanced) (see Peng et al., 2021 for a review of the test). The examinees took the HSK Level 4 written test (score range: 0–300) and the intermediate-level speaking test (score range: 0–100). The mean of the combined test scores was 229.03 ($SD = 51.84$, range: 142.25–328.75), suggesting that the examinees had roughly intermediate-mid to advanced-mid level of proficiency.

3.2 Instrument

All examinees responded to a 12-item computerized oral DCT consisting of six speech act items representing request ($k = 2$), refusal ($k = 2$), and compliment response ($k = 2$), as well as six pragmatic routine items. The Appendix shows a list of these scenarios. These items came from a larger project assessing pragmatic development in L2 Chinese (e.g., Li et al., 2019; Taguchi et al., 2016; Xiao et al.,

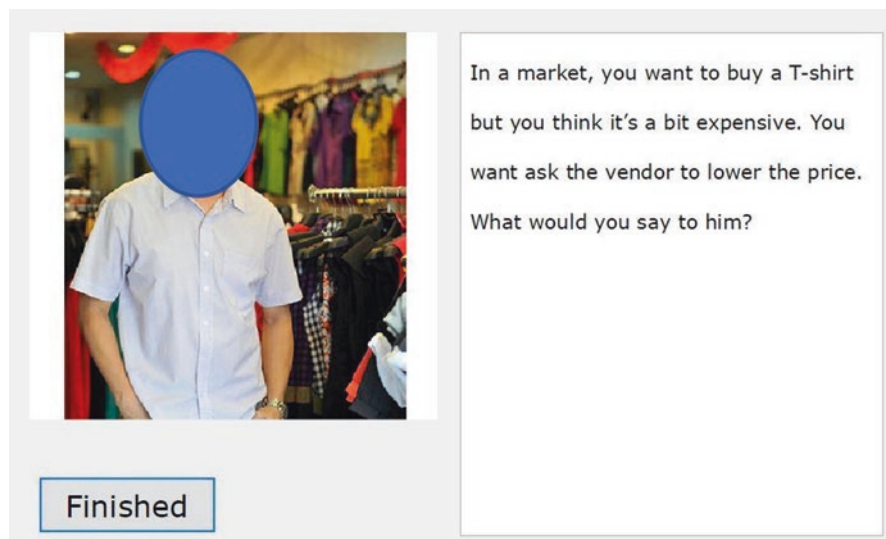


Fig. 1 A screenshot of an Oral DCT item

2019). In responding to the oral DCT, examinees first heard a scenario description in English and at the same time saw a picture illustrating the scene. After the audio was done, a beep reminded the examinees to start saying what they would say in that scenario. There was no time limit for the assessment items. Examinees' oral responses were recorded in the computer, and the audio files were evaluated by a group of non-expert raters (see above Fig. 1).

3.3 Non-expert Raters

A total of 101 non-expert raters were recruited from a major south-eastern city in the US. They were all native Chinese speakers coming from Mainland China and were enrolled in colleges and universities at the time of this study. The mean length of stay in the U.S. was 30.58 months ($SD = 32.68$). There were 64 males and 47 females, with a mean age 24.19 years ($SD = 4.65$ years). None of the raters were in the fields of linguistics or applied linguistics, and all reported no or highly limited experience of interacting with learners of L2 Chinese.

3.4 Rating Procedures

The oral responses of the 51 examinees were assigned to 15 batches. Each batch contained the data of nine examinees, including the same six examinees shared across all batches (which served as anchors for linking different raters' performance

when performing the Rasch analysis) and three different, randomly selected examinees. The 101 raters were randomly assigned to evaluate one of the 15 batches of examinee data; each batch of examinee data was evaluated by seven, eight or nine raters. They saw the same DCT scenarios as the examinees did, listened to the examinees' oral productions one by one, and performed a Yes/No binary judgment on whether an oral response fulfilled the communicative goal as required by each scenario. The raters were also encouraged, but not required, to verbalize their scoring rationale in Chinese after making each judgment. No specific guidelines were given to the raters and they were free to comment on any aspect of the oral responses. The raters' verbal protocols were recorded and later transcribed for analysis.

Before starting their judgment and verbalization task, the raters received a brief warmup exercise, during which they tried several practice items and were familiarized with the verbalization procedure. Data collection was conducted individually for each rater in a quiet room on campus.

3.5 Data Analyses

To answer RQ 1, we collected a total of 10,908 binary judgments (101 raters \times 9 examinees for each rater \times 12 DCT items for each examinee). All "Yes" judgments were converted to the score of "1" and all "No" judgments the score of "0." Due to two missing data points, the total number of judgments for statistical analysis was 10,906. Out of the 10,906 judgments, 8289 (or 76%) were "Yes", and the remaining 2617 (or 24%) were "No". We built a three-facet Rasch dichotomous model including raters ($n = 101$), examinees ($n = 51$), and oral DCT items ($k = 12$). The quantitative analysis was performed with the software FACETS Version 3.71.3.

To answer RQ 2, due to an unexpected loss of a portion of the verbal protocol data, analysis was based on the data from 81 raters who evaluated 48 examinees. Out of the 8748 potential verbal protocols (i.e., 81 raters \times 9 examinees \times 12 scenarios), the 81 raters provided 2753 verbal protocols (a 31.47% response rate). Based on these protocols, the three researchers of this study followed a data-driven approach (Youn, 2015) to developing our coding scheme. This involved a bottom-up, iterative procedure by reviewing all verbal protocols in order to extract and refine our codes and the entire coding scheme. The finalized coding scheme included 16 first-order codes (rating criteria), which were grouped into three major categories: *holistic meaning expression*, *linguistic expressions*, and *interaction*. The total instances of coding were 2945. The three researchers went through and discussed all instances of coding together to reach consensus. Following is the coding scheme that provides definitions of each first-order code (rating criteria) with representative examples from our data. Due to space limit, only English translations are provided for the examples. All coding was performed through NVivo Version 12.

1. Holistic meaning expression

1.1 Comprehensibility of meaning

Definition: Overall comprehensibility of an utterance, ease of understanding the speaker's intention

Example #1, for Scenario #1 (Cashier) (see scenario description in the [Appendix](#))

He has already expressed that he wants to buy a jacket, therefore I could understand it clearly when he was asking where he is going to pay.

1.2 Incomprehensibility of meaning

Definition: Overall incomprehensibility of meaning, difficulty in understanding the speaker's intention

Example #2, for Scenario #7 (Wrong phone call)

I only understood the part "I am" and didn't understand what he said afterwards.

1.3 Misunderstanding

Definition: an utterance that may cause misunderstanding.

Example #3, for Scenario #5 (Bargain)

What she said was "this T-shirt is too expensive, and I don't have money", which makes people think that she might not mean to ask the peddler to lower the price; instead, she might not want to buy this T-shirt. Just a little bit like, her expression could cause the peddler to misunderstand what she means and therefore is not willing to continue the conversation with her anymore.

1.4 Incomplete meaning

Definition: an utterance that does not fully express the intended meaning by leaving out important information (i.e., lacking semantic formula)

Example #4, for Scenario #3 (Presentation)

He didn't express his refusal; he just said that he was sorry, which could be counted as a half refusal, but he did not provide any reasons.

2. Linguistic expressions

2.1 Code switching

Definition: an utterance that includes the use of English words/phrases

Example #5, for Scenario #8 (Restaurant)

Restaurant waiters, with their English proficiency, won't be able to understand the meaning of "carry away".

2.2 Word choice

Definition: an utterance that includes wrongly used word(s)

Example #6, for Scenario #6 (Photo)

She chose the wrong verb and said "to make a photo", which expresses a completely different meaning as for "taking photos".

2.3 Key expression

Definition: production (or lack thereof) of keyword(s) that renders success (or lack thereof) in meaning expression

Example #7, for Scenario #1 (Cashier)

He mentioned the keywords “to pay”, and the salesperson should be able to understand what he meant.

2.4 Incomplete utterance

Definition: an utterance that is syntactically incomplete due to a lack of linguistic knowledge

Example #8, for Scenario #8 (Restaurant)

Because he didn’t know how to express the meaning “taking the food away”. He just came up directly and asked the waiter, but he didn’t know what to say next.

2.5 Grammar

Definition: syntactic or morphosyntactic features of an utterance that may interfere with or enhance meaning expression

Example #9, for Scenario #6 (Photo)

There are some problems with his word order. The adverbial is not put in the correct position, and there is no preposition in the sentence; but we could understand him in communication.

2.6 Pronunciation

Definition: clarity and accuracy of pronunciation that may interfere with or enhance meaning expression

Example #10, for Scenario #1 (Cashier)

First of all, her pronunciation is not accurate. She said that “I want to sell this”, “where I can buy it”. If I were the clerk, I would ask what you want to sell, and what you meant by saying where to buy. I don’t understand what you are talking about.

2.7 Intonation

Definition: intonational features that may interfere with or enhance meaning expression and/or politeness

Example #11, for Scenario #4 (Essay)

“Do you think it is very interesting?”. It may not sound very polite to use a rhetorical question in Chinese.

2.8 Fluency

Definition: temporal features that may interfere with or enhance meaning expression

Example #12, for Scenario #10 (Cell phone)

Because she speaks intermittently, I couldn’t hear what she was talking about.

2.9 Nativelikeness

Definition: an utterance or expression that may or may not conform to native speakers’ intuition

Example #13, for Scenario #5 (Bargain)

What he wants to express should be that the T-shirt is a little expensive, wishing it to be cheaper. He said, “Why is it so expensive”, meaning that “this is a little expensive”. However, the way he said it is quite different from what we are used to, and we might not be able to understand what he wants to express the moment we hear it in actual communication.

2.10 Politeness

Definition: level of politeness that may interfere with or enhance effectiveness in interaction

Example #14, for Scenario #4 (Essay)

Professor Xiao praised his travel essay, but he said “thank you, do you understand this?” Although he might not intend to offend Professor Xiao, this expression is very offensive to the interlocutor.

3. Interaction

3.1 Turn management

Definition: an utterance that is perceived to connect well or poorly with prior or subsequent turns

Example #15, for Scenario #4 (Essay)

He did not respond to the interlocutor’s comments, what he said was a totally different thing from what the interlocutor had said. The interlocutor already asked to discuss this essay with him and expressed that this essay is interesting. But he is still asking “Can I discuss this essay with you?”

3.2 Contextualization

Definition: visualization specific context of communication

Example #16, for Scenario #5 (Bargain)

He expresses that it is very expensive while holding a T-shirt, therefore the peddler should understand that he wants to buy this T-shirt at a cheaper price.

4 Results

4.1 RQ1. Non-expert Raters' Scoring Behavior

RQ1 focused on non-expert raters' behavior of scoring L2 pragmatic performance. Figure 2 is an output graph of the Rasch model. The first column on the left represents the logit scale, on which rater severity, examinee ability, and item difficulty are measured. The second column shows the distribution of rater severity with each asterisk (*) representing two raters and each dot (.) one rater. Harsher raters appear in higher positions than more lenient raters. The third column displays the ability distribution of the 51 examinees. A higher position on the logit scale corresponds to a higher ability level, and vice versa. The last column indicates the distribution of items in terms of difficulty level. More difficult items occupy higher positions on the scale than easier items (e.g., Item #1, Cashier, was the most difficult one).

Rasch calibrated statistics showed that rater severity measures spread across 3.41 logits (i.e., from 1.76 to -1.65 logits), indicating variability in scoring severity among the raters. Indeed, the corresponding rater separation index was 1.92 (or 2.90 strata) with a reliability coefficient of .79, meaning that the raters can be grouped according to three statistically distinct levels of severity. Importantly, all (100%)

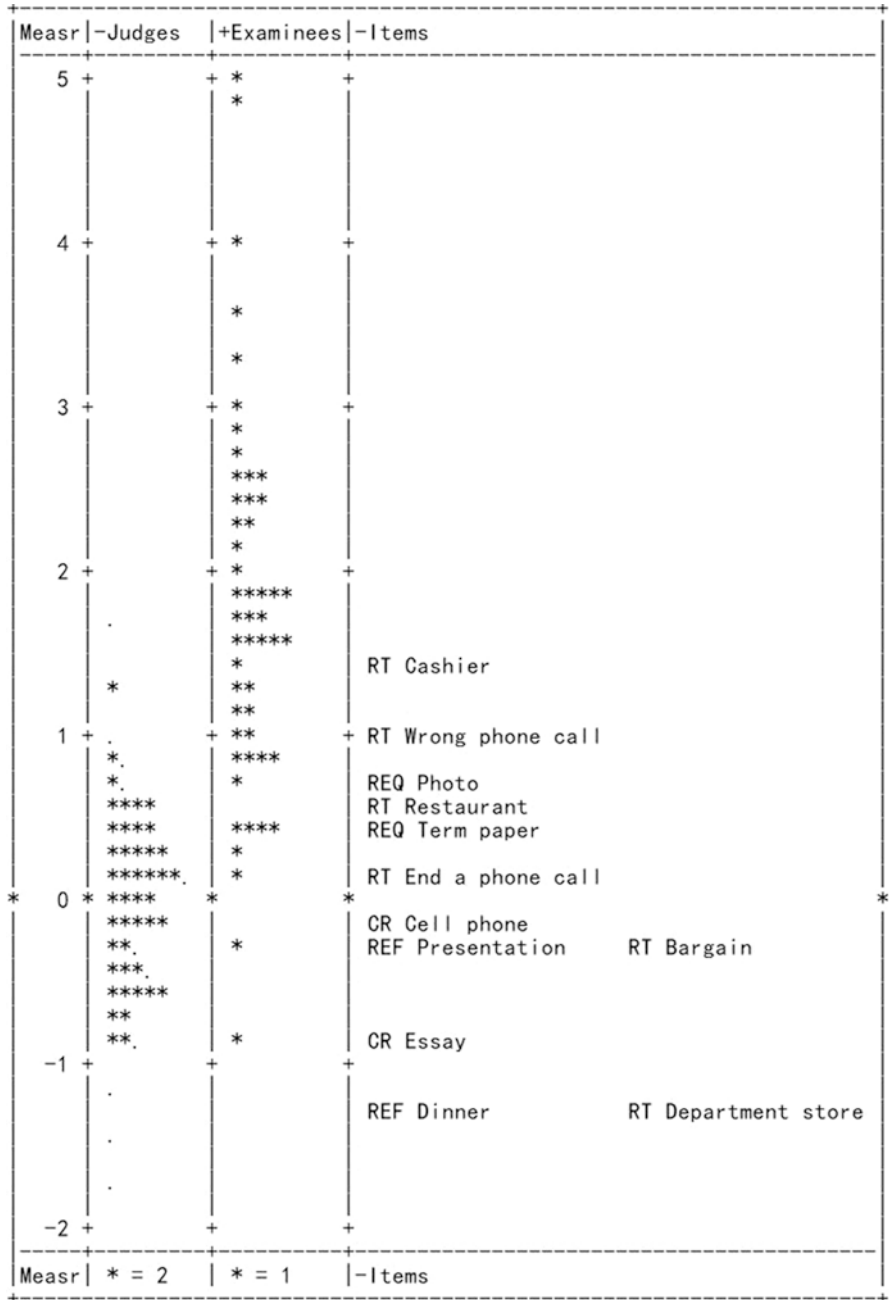


Fig. 2 Wright map. (Note. CR Compliment response, REQ Request, REF Refusal, RT Routine)

raters' infit MnSq values fell within the 0.7–1.3 range, suggesting satisfactory model fit, i.e., all raters' judgments conformed to the expectations of the Rasch model. This means that the raters' judgments were consistent. It is also relevant to briefly report the examinee statistics. Examinee ability measures spanned across 6.08 logits (from 5.27 to -0.81 logits), with an average of 1.74 logits ($SD = 1.19$). The examinee separation index was 2.50 (or 3.67 strata) with a reliability coefficient of .86. This means that the examinees could be reliably grouped into more than three distinct ability levels. Moreover, the infit MnSq statistics of 50 out of the 51 examinees (or 98%) fell within the 0.7–1.3 range, suggesting satisfactory model fit of individual examinees' item responses. Turning to the item statistics, the item difficulty measures spread by 2.82 logits (from 1.47 to -1.35 logits).

We further conducted two sets of bias/interaction analyses to examine: (1) whether the raters were more or less severe in scoring individual examinees, and (2) whether they were more or less severe in scoring according to individual items. For the rater \times examinee bias/interaction analysis, the purpose was to test the null hypothesis that “there is no statistically discernible bias in each rater's ratings towards individual examinees.” Out of 906 bias/interaction terms, only one (or 0.11%) reached statistical significance. This is substantially below the commonly accepted 5% misfit ratio. For the rater \times item bias/interaction analysis, we were interested in testing the null hypothesis that “there is no statistically discernible bias in each rater's ratings towards individual items”. Out of 1212 interaction terms, 25 (or 2.06%) were statistically significant, which is also below the commonly accepted 5% threshold.

In summary, the non-expert raters' scoring behavior met the expectations of the Rasch model in terms of scoring consistency, yet the raters varied significantly in scoring severity. The raters as a group showed very limited instances of bias in scoring towards individual examinees and/or according to assessment items.

4.2 RQ2. Non-expert Raters' Rating Criteria

RQ 2 examined the criteria that our non-expert raters drew on to evaluate task fulfillment of different speech acts and pragmatic routines. Table 1 displays the frequencies of all first-order codes (i.e., criteria) grouped according to three major categories (i.e., *holistic meaning expression*, *linguistic expressions*, and *interaction*) and for speech acts and routines, respectively. In presenting the findings, we will refer to the examples in the coding scheme (see the Method section).

The percentage statistics in Table 1 show similarities and differences in raters' criteria for assessing speech acts and pragmatic routines. Regarding similarities, *holistic meaning expression* was the most frequently referenced among the three major categories, accounting for 55.97% of the total instances of codes for speech acts and 55.06% for routines. Raters often commented holistically on whether an utterance's meaning was comprehensible (i.e., *comprehensibility of meaning*, see Example #1 in the coding scheme) or incomprehensible (i.e., *incomprehensibility of*

Table 1 Distribution of first-order codes (criteria) between speech acts and pragmatic routines

Categories	First-order codes (criteria)	Speech acts		Routines	
Holistic meaning expression	Misunderstanding	86	6.22%	55	3.52%
	Incomplete meaning	135	9.76%	90	5.76%
	Comprehensibility of meaning	397	28.71%	442	28.30%
	Incomprehensibility of meaning	156	11.28%	273	17.48%
	<i>Subtotal</i>	774	55.97%	860	55.06%
Linguistic expressions	Code switching	62	4.48%	133	8.51%
	Word choice	27	1.95%	57	3.65%
	Key expression	49	3.54%	181	11.59%
	Incomplete utterance	16	1.16%	43	2.75%
	Grammar	25	1.81%	12	0.77%
	Pronunciation	72	5.21%	103	6.59%
	Intonation	4	0.29%	4	0.26%
	Fluency	29	2.10%	28	1.79%
	Nativeness	10	0.72%	16	1.02%
	Politeness	19	1.37%	15	0.96%
	<i>Subtotal</i>	313	22.63%	592	37.90%
Interaction	Turn management	272	19.67%	32	2.05%
	Contextualization	16	1.16%	47	3.01%
	<i>Subtotal</i>	288	22.82%	79	5.06%
Uncoded		8	0.58%	31	1.98%
Total		1383	100.00%	1562	100.00%

meaning, see Example #2). To a far lesser extent, raters also based their judgments on whether an utterance might lead to *misunderstanding* (see Example #3) and whether an utterance fully expressed the intended communicative function expected in a specific scenario (i.e., *incomplete meaning*, Example #4).

On the other hand, the non-expert raters differentially drew on the other two larger categories of criteria according to the targeted pragmatic features. As Table 1 shows, raters commented on aspects of *linguistic expressions* more frequently when assessing pragmatic routines (36.94%) than speech acts (21.26%). A closer examination of the individual criteria within this category revealed a nuanced picture. To begin with, the differences between speech acts and routines mainly came from five criteria, and three of these criteria were about vocabulary knowledge: *code switching* (Example #5), *word choice* (Example #6), and *key expression* (Example #7). Table 1 shows that raters referred to *code switching* and *word choice* nearly twice as frequently in assessing routines as in assessing speech acts; the difference in *key expression* was even larger. Another criterion that was used with higher frequency in assessing routines than speech acts was *incomplete utterance* (Example #8). Still another criterion with notable difference between routines and speech acts was *grammar* (Example #9); but this time the frequency was higher for speech acts than for routines. Different from the previous five criteria, the raters showed little difference between routines and speech acts for the following criteria: *pronunciation*

(Example #10), *intonation* (Example #11), *fluency* (Example #12), *nativelikeness* (Example #13), and *politeness* (Example #14). Among these criteria, raters commented on *pronunciation* more frequently than the other criteria.

In terms of the larger category *interaction*, it carried heavier weight for assessing speech acts (22.20%) than for routines (6.02%). Among the three criteria within this category, *turn management* (Example #15) is where there was a major gap between routines (2.05%) and speech acts (19.67%). As it turned out, out of the 272 references to *turn management* under speech acts, compliment response accounted for 94.85%, whereas refusal and request took 5.15% and 0%, respectively. For the remaining criterion *contextualization* (Example #16), raters referred to it more than twice as frequently for routines as for speech acts.

In summary, our non-expert raters relied on three major categories of criteria to evaluate fulfillment of pragmatics tasks involving speech acts and routines. They predominantly focused on the criteria under the larger category *holistic meaning expression* for evaluating both speech acts and pragmatic routines. They appeared to prioritize the criteria under the larger category of *linguistic expressions* when scoring pragmatic routines; meanwhile, they paid more attention to the criteria under the larger category of *interaction* when evaluating speech acts.

5 Discussion

RQ 1 focused on non-expert raters' scoring behavior. The raters varied significantly in scoring severity, yet their scoring performances were highly consistent. Moreover, there were only very limited instances of scoring bias towards individual examinees or items. These findings echo existing research on the scoring behavior of expert raters in assessing L2 pragmatics (e.g., Liu & Xie, 2014; Youn, 2007). Different from previous studies where expert raters were given predetermined rating criteria and received training on scoring, the non-expert raters in this study were not given any uniform, *a priori*, assessment criteria, nor did they receive training on scoring pragmatic performance. Instead, our non-expert raters were free to utilize their own criteria to judge examinees' fulfillment of the pragmatics tasks. In previous studies, expert raters were typically asked to score pragmatic performance according to multiple score bands, which is arguably more cognitively complicated than the binary judgments that our non-expert raters did in this study. Because the binary judgments were relatively straightforward, the raters probably did not need specialized knowledge or training, and could instead rely on their native-speaker intuitions to make judgments. Hence, it was likely that the straightforwardness of the judgment task contributed to the high level of scoring consistency in this study. It would be interesting to examine non-expert raters' scoring consistency based on a rating scale with multiple score bands.

Severity in scoring, on the other hand, showed considerable variation among the 101 non-expert raters, with the rater severity measures spanning across 3.41 logits with a separation index of 1.92 (or 2.90 strata). Because each rater scored only a

subset of the examinees, large individual differences in scoring severity tended to have a major impact on examinee scores. In L2 performance assessment, a typical threshold is that the range of examinee ability is roughly twice (or more) as wide as the range of rater severity; when this threshold is met, the impact on individual raters' scoring severity on examinee test scores is considered as acceptable (Myford & Wolfe, 2000). In pragmatics research focusing on rater behavior, only a few studies reported both rater and examinee statistics, and there are variations across studies in meeting this criterion. For example, Youn (2007) reported a rater severity range of 0.52 logits and an examinee ability range of 0.51 logits, which is way below the threshold; yet Li et al. (2019) found a rater severity range of 0.56 logits and an examinee ability range of 3.75 logits, which is clearly above the threshold. In this study, the examinees' ability range was 6.08 logits, which is nearly twice the range of rater severity (i.e., 3.41 logits). We suspect that a lack of rater training (which was intentional in this study) and raters' personality attributes (i.e., being harsher or more lenient) may have resulted in the variability in rater severity in this study. Our non-expert raters, unlike the expert raters in previous studies (e.g., Li et al., 2019; Youn, 2007), did not have an opportunity to discuss and calibrate their scoring criteria as a group (and it was logistically impractical to do so given the large number of raters recruited). It would be interesting to examine the extent to which non-expert raters' scoring severity can be homogenized by introducing appropriate rater training sessions, which may help ameliorate the influence of personality traits on scoring severity.

RQ 2 examined the non-expert raters' cognitive processes during scoring, focusing on the similarities and differences as they evaluated two different types of pragmatic features, i.e., speech acts and routines. Verbal protocol analysis showed that the raters predominantly oriented towards criteria related to *holistic meaning expression* regardless of pragmatic features. This finding makes sense because the raters were instructed to judge task fulfillment, i.e., whether the intended meaning was conveyed in a specific scenario, which clearly depends on the success in conveying the intended meaning.

Our non-expert raters also paid attention to various criteria under the larger categories of *linguistic expressions* and *interaction*, where there were notable differences between speech acts and pragmatic routines. While Li et al. (2019) demonstrated that individual raters' scoring behavior varied according to different pragmatic features, the results to be discussed here complement their findings by uncovering how raters' underlying cognitive processes may vary based on different pragmatic features. Specifically, under *linguistic expressions*, raters commented on vocabulary knowledge (i.e., criteria of *key expression*, *code switching*, and *word choice*) more frequently for assessing routines than speech acts; the pattern was revised for the criterion *grammar*, which was cited more frequently for evaluating speech acts than routines. These differences likely reflect the unique characteristics of the two pragmatic features and echo existing findings on the acquisition of speech acts and routines (discussed below).

To begin with, speech acts such as requests and refusals typically entail the coordination of various semantic formulae (e.g., providing justifications, thanking, and the focal request/refusal expression *per se*) and the production of syntactically

complex forms. For example, the examinee expression associated with Example #9 was 你可以拍照片我吗?(*Could you take my picture?*). This grammatically incorrect utterance that involves the question structure 可以...吗? (*Could ... question particle?*) lacks a complex preposition structure, as the rater pointed out in Example #9. In contrast, pragmatic routines, being fixed or semi-fixed linguistic expressions, are syntactically simpler and semantically less complicated than speech acts. This means that each word in a routine expression plays an important role; oftentimes, one keyword or one short expression could determine the success or failure of producing a pragmatic routine, as Examples #5 and #7 can show. In addition, previous studies on the acquisition of Chinese pragmatic routines reported that an important strategy that learners employed to develop their ability to produce routines was to use core lexical items (e.g., Bardovi-Harlig & Su, 2018; Li et al., *in press*; Taguchi et al., 2013). Regarding speech acts, researchers have reported that L2 Chinese learners often experienced difficulty in incorporating morphosyntactic and lexical devices into request utterances (e.g., Li, 2014; Wen, 2014), and that learners gradually developed the ability to produce more semantically sophisticated refusals (e.g., Tang et al., 2021). In this study, the non-expert raters were able to intuitively adjust their evaluation criteria and orient to different aspects of examinees' performance according to speech acts and pragmatic routines, and they did so without knowledge of relevant pragmatics theories and/or research findings.

Under the larger category of *linguistic expressions*, our non-expert raters also paid attention to *pronunciation, intonation, fluency, politeness, and nativelikeness*. These results corroborate prior research on what expert raters focus on when evaluating speech acts in L2 English (Alemi & Tajeddin, 2013; Liu & Xie, 2014; Sydorenko et al., 2014; Taguchi, 2011; Walters, 2007). While existing studies typically featured only a small number of expert raters, the relatively large number of non-expert raters in this study, along with its focus on Chinese as the target language and on two types of pragmatic features, can add to the generalizability of existing research findings. It is encouraging to know that non-expert raters are able to orient to aspects of L2 pragmatic performance just like expert raters do.

The third larger category that emerged from our protocol data was *interaction*, where there were also considerable differences between speech acts and pragmatic routines. There was a large gap in frequency of reference regarding *turn management* (i.e., 19.67% for speech acts and 2.05% for routines). The relatively frequent comments on *turn management* for speech acts was a bit surprising at first glance, because this study adopted a single-turn oral DCT, which did not allow turn taking or meaning negotiation. In hindsight, this finding was likely due to the characteristics of the speech acts under investigation in this study. In particular, the scenarios involving compliment responses and refusals, by nature, involved a responding turn rather than an initiating turn (which was the case for the request scenarios). Clearly, *turn management* is a key skill in scenarios involving compliment responses and refusals because task fulfilment depends on examinees' ability to produce a turn that connects naturally and sensibly to the previous turn, as Example #15 can show. Indeed, 94.85% of our raters' comments on turn management were found in the compliment response scenarios and 5.15% in the refusal scenarios. On the other hand, none of the routine scenarios necessitate a responding turn, which could

explain why *turn management* carried much lighter weights in our raters' evaluation of routines.

The last criterion where there was a notable difference between speech acts and routines was *contextualization* (under the larger category of *interaction*). Raters cited this criterion more than twice as frequently in assessing routines (3.01%) as in assessing speech acts (1.16%). As Example #16 can show, even though the scenario description and the accompanying photo (Fig. 1) do not indicate or show a speaker holding a T-shirt in hand, trying to bargain a sales price with the street vendor, the rater was able to mentally visualize the specific scene by drawing on personal experiences and/or observations. Comparatively speaking, *contextualization* was more prominent among routine scenarios than speech act scenarios. This is likely because routines, by definition, are tied to specific contexts of communication, i.e., there is a relatively fixed connection between a routine expression and a particular scenario – Kecskes (2016) even coined the term *situationally bound utterances* to refer to pragmatic routines. In comparison, pragmalinguistic forms of speech acts can often be used across different scenarios, thus the connection between pragmalinguistic form and context is weaker than that between routines and context. It would therefore be easier for raters to visualize a specific scene for routines than for speech acts.

6 Conclusions and Boundary Crossing Implications

In crossing the boundaries regarding rater eligibility, target construct, and target language, as identified in the literature review section, this study represented an initial effort to examine non-expert raters' scoring behavior and cognition involved in assessing pragmatics in L2 Chinese. Concerning scoring behavior, despite considerable variability in judgment severity, the non-expert raters performed scoring consistently, with very limited instances of scoring biases. Concerning rater cognition, the raters were primarily oriented to *holistic meaning expression* in judging examinees' task fulfillment regardless of pragmatic features. However, they focused more on criteria related to *linguistic expressions* (notably those related to vocabulary knowledge) in evaluating pragmatic routines than speech acts, and more on criteria related to *interaction* (notably the criterion of *turn management*) when assessing speech acts than routines. Such variability in rater cognition according to targeted pragmatic features can be explained by the characteristics of speech acts and pragmatic routines.

By crossing multiple boundaries, this study can have practical implications for pragmatics assessment and L2 teaching in general. In crossing the rater eligibility boundary by focusing on non-expert raters (in contrast to previous studies' predominant focus on expert raters), this study demonstrates that untrained native speakers (of Chinese), as important stakeholders of pragmatics assessment, are actually able to evaluate L2 pragmatic performance and achieve satisfactory scoring quality, provided that the scoring task is relatively straightforward and that the stakes of the intended pragmatics assessment are relatively low. Including

non-expert native speaker raters in the process of pragmatics assessment should enhance the validity of score interpretation in the assessment context of the target language community, because these are the people that L2 learners are most likely to interact with. Furthermore, in light of the potential of Chinese becoming a lingua franca (Gil, 2021), it would be worthwhile to further cross the rater eligibility boundary by investigating the feasibility of including non-expert, non-native speakers of Chinese who use the language for their professions.

Moreover, in crossing the target construct boundary by including pragmatic routines in addition to speech acts, results of this study indicate that non-expert raters adjust evaluation criteria according to different pragmatic features. Hence, if a goal of L2 pragmatics assessment is to inform examinees of their strength and weakness when they interact with potential interlocutors, it would be important to develop evaluation criteria according to targeted pragmatic features and, perhaps also adjust the weights of such criteria accordingly. These issues would not have surfaced in this study, and would not inform future studies, if we or L2 pragmatics researchers alike were limited by the target construct boundary and focused predominantly on speech acts. Future research can continue to cross the target construct boundary by including more varied pragmatic features in investigating rater cognition and scoring behavior.

Finally, implications of our findings can also cross the boundary of pragmatics assessment to inform L2 (Chinese) teaching in general. The fact that our non-expert raters paid predominant attention to criteria under *holistic meaning expression* highlights the importance of focusing on communicative function (i.e., expressing intended meaning) in L2 instruction. While formal aspects of linguistic expressions do matter in the evaluation of task fulfillment, vocabulary knowledge and, to a lesser extent, pronunciation skills appear to be more important than grammatical knowledge based on our rater protocol analysis. While Chinese language instructors often tend to emphasize grammatical structures in instruction, our findings suggest that grammatical accuracy may only play a very minor role in determining the success of getting one's message across. Moreover, skills such as turn management, which is often not emphasized in (Chinese) language classrooms or in textbooks, should be highlighted to various degrees according to instructional targets (e.g., speech acts vs. pragmatic routines).

Appendix: List of 12 Scenarios

Item numbers indicate order of appearance in the Oral DCT.

Speech act scenarios

Compliment response items

#4 (Essay) You wrote an essay about your travel experience and submitted it to Professor Xiao's class. Today, you meet him in the hallway and you start to

talk to each other. During your conversation, Professor Xiao says: “Oh, by the way, I read your essay and it is really interesting.” What would you say to him?

#10 (Cell phone) You meet your friend Xiao Wang in the hallway. Xiao Wang sees your newly purchased cell phone and says: “Is this your new cell phone? It looks really fancy!” How would you respond to Xiao Wang?

Refusal items

#2 (Dinner) You meet your friend Xiao Li after class. Xiao Li invites you to dinner with his friend but you don’t want to go. What would you say to Xiao Li?

#3 (Presentation) You come to Professor Li’s office to ask a few questions. Before you leave, she asks you to do your presentation one week earlier than you originally scheduled. However, you don’t want to do that. What would you say to Professor Li?

Request items

#6 (Photo) You meet your friend Xiao Li at a party today. You want to ask Xiao Li to take your picture. What would you say to him?

#12 (Term Paper) Today is the deadline for submitting your term paper, but you don’t have it finished because you were sick. So you want to ask Professor Sun for an extension. Now you come to Professor Sun’s office. What would you say to him?

Pragmatic routine scenarios

#1 (Cashier) At a department store, you cannot find where the cashier is. You want to ask this shop assistant for this. How would you ask him?

#5 (Bargain) In a market, you want to buy a T-shirt but you think it’s a bit expensive. You want to ask the vendor to lower the price. What would you say to him?

#7 (Wrong phone call) When you answer your phone, you hear a young man’s voice. Obviously, he dialed your number by mistake. What would you say to him?

#8 (Restaurant) In a restaurant, you want to take the leftovers with you. What would you say to this waitress?

#9 (End a phone call) You and your friend are talking on the phone. It seems that you both have said all you want to say, so you would like to end your conversation. What would you say to her?

#11 (Department store) In a department store, a shop assistant asks whether you would like to buy anything. You do not intend to buy anything. What would you say to her?

References

- Alemi, M., & Tajeddin, Z. (2013). Pragmatic rating of L2 refusal: Criteria of native and nonnative English teachers. *TESL Canada Journal*, 30(7), 63–81.
- Bardovi-Harlig, K. (2003). Understanding the role of grammar in the acquisition of L2 pragmatics. In A. Martinez-Flor, E. U. Juan, & A. F. Guerra (Eds.), *Pragmatic competence and foreign language teaching* (pp. 25–44). Universitat Jaume I.
- Bardovi-Harlig, K., & Su, Y. (2018). The acquisition of conventional expressions as a pragmatic-linguistic resource in Chinese as a foreign language. *The Modern Language Journal*, 102(4), 732–757.
- Chen, Y., & Liu, J. (2016). Constructing a scale to assess L2 written speech act performance: WDCT and e-mail tasks. *Language Assessment Quarterly*, 13(3), 231–250.
- Gil, J. (2021). *The rise of Chinese as a global language: Prospects and obstacles*. Palgrave Macmillan.
- Grabowski, K. (2013). Investigating the construct validity of a role-play test designed to measure grammatical and pragmatic knowledge at multiple proficiency levels. In S. Ross & G. Kasper (Eds.), *Assessing second language pragmatics* (pp. 149–171). Palgrave Macmillan.
- Han, Q. (2016). Rater cognition in L2 speaking assessment: A review of the literature. *Working Papers in TESOL & Applied Linguistics*, 16(1), 1–24.
- Hudson, T., Detmer, E., & Brown, J. D. (1995). *Developing prototypic measures of cross-cultural pragmatics (technical report #7)*. University of Hawaii, Second Language Teaching & Curriculum Center.
- Kasper, G. (2006). Speech acts in interaction: Towards discursive pragmatics. In K. Bardovi-Harlig, J. C. Felix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics and language learning: Vol. 11* (pp. 281–314). University of Hawai'i at Manoa: National Foreign Language Resource Center.
- Kecskes, I. (2016). Situation-bound utterances in Chinese. *East Asian Pragmatics*, 1(1), 108–126.
- Leech, G. (1983). *Principles of pragmatics*. Longman.
- Li, S. (2014). The effects of different levels of linguistic proficiency on the development of L2 Chinese request production during study abroad. *System*, 45, 103–116.
- Li, S. (2021). Pragmatics assessment in English as an international language (EIL). In Z. Tajeddin & M. Alemi (Eds.), *English as an international language: Pragmatic pedagogy* (pp. 191–211). Routledge.
- Li, S., Taguchi, N., & Xiao, F. (2019). Variations in rating scale functioning in assessing pragmatic performance in L2 Chinese. *Language Assessment Quarterly*, 16(3), 271–293.
- Li, S., Taguchi, N., & Xiao, F. (in press). Effects of proficiency on the development of pragmatic routine production in L2 Chinese. In F. Xiao (Ed.), *Second language Chinese development: A longitudinal perspective*. Lexington Books.
- Liu, J. (2006). *Measuring interlanguage pragmatic knowledge of Chinese EFL learners*. Peter Lang.
- Liu, J., & Xie, L. (2014). Examining rater effects in a WDCT pragmatics test. *Iranian Journal of Language Testing*, 4(1), 50–65.
- McNamara, T., Knoch, U., & Fan, J. (2019). *Fairness, justice, and language assessment*. Oxford University Press.
- Myford, C., & Wolfe, E. (2000). *Monitoring sources of variability within the test of spoken English assessment system* (Vol. 2000, pp. i–51). Educational Testing Service.
- Peng, Y., Yan, W., & Cheng, L. (2021). Hanyu Shuiping Kaoshi (HSK): A multi-level, multi-purpose proficiency test. *Language Testing*, 38(2), 326–337.
- Roever, C. (2007). DIF in the assessment of second language pragmatics. *Language Assessment Quarterly*, 4(2), 165–189.
- Sydorenko, T., Maynard, C., & Guntly, E. (2014). Rater behaviour when judging language learners' pragmatic appropriateness in extended discourse. *TESL Canada Journal*, 32(1), 19–41.
- Taguchi, N. (2011). Rater variation in the assessment of speech acts. *Pragmatics*, 21(3), 453–471.

- Taguchi, N. (2012). *Context, individual differences, and pragmatic competence*. Multilingual Matters.
- Taguchi, N. (2019). Second language acquisition and pragmatics: An overview. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 1–14). Routledge.
- Taguchi, N., & Li, S. (2021). Contrastive pragmatics and second language pragmatics: Approaches to assessing L2 speech act production. *Contrastive Pragmatics*, 2(1), 1–23.
- Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. Oxford University Press.
- Taguchi, N., Li, S., & Xiao, F. (2013). Production of formulaic expressions in L2 Chinese: A developmental investigation in a study-abroad context. *Chinese as a Second Language Research*, 2(1), 23–58.
- Taguchi, N., Xiao, F. & Li, S. (2016). Development of pragmatic knowledge in L2 Chinese: Effects of intercultural competence and social contact on speech act production in a study abroad context. *The Modern Language Journal*, 100(4), 775–796.
- Tang, X., Taguchi, N., & Li, S. (2021). Social contact and speech act strategies in a Chinese study abroad context. *Study Abroad Research in Second Language Acquisition and International Education*, 6(1), 3–31.
- Thomas, J. (1983). Cross-cultural pragmatic failure. *Applied Linguistics*, 4, 91–111.
- Timpe, V. (2013). *Assessing intercultural communicative competence*. Peter Lang.
- Walters, S. F. (2007). A conversation-analytic hermeneutic rating protocol to assess L2 oral pragmatic competence. *Language Testing*, 24(2), 155–183.
- Wen, X. (2014). Pragmatic development: An exploratory study of requests by learners of Chinese. In Z. Han (Ed.), *Studies in second language acquisition of Chinese* (pp. 30–56). Multilingual Matters.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370–371.
- Xiao, F., Taguchi, N., & Li, S. (2019). Effects of proficiency sub-skills on pragmatic development in L2 Chinese study abroad. *Studies in Second Language Acquisition*, 41(2), 469–483.
- Youn, S. (2007). Rater bias in assessing the pragmatics of KFL learners using facets analysis. *Second Language Studies*, 26(1), 85–163.
- Youn, S. (2015). Validity argument for assessing L2 pragmatics in interaction using mixed methods. *Language Testing*, 32(2), 199–225.
- Young, R. (2011). Interactional competence in language learning, teaching, and testing. In E. Hinkel (Ed.), *Handbook of research in language learning and teaching* (pp. 426–443). Routledge.

Early Home and Community Support in Later Chinese Heritage Language Literacy Development



Haomin Zhang, Xi Cheng, and Jiexin Lin

Abstract Prior language and literacy support from home, community and heritage language (HL) schools provides young HL learners with linguistic and metalinguistic foundations for language and literacy development (Chinen & Tucker *Heritage Language Journal* 3:27-59, 2005; Koda et al. Chinese as a heritage language: Fostering rooted world citizenry (pp. 125–135). University of Hawai'i, National Foreign Language Resource Center. 2008; Mori & Calder *Foreign Language Annals* 46:290-310, 2013). To date, few studies have explored the contribution of early language experiences to later HL literacy development. The current study aimed to investigate the role of learner-external and input-based language and literacy support in adult HL literacy development. Two hundred and nine Chinese as a heritage language (CHL) learners participated in the study. They completed a language and literacy background survey as well as a series of Chinese reading measurements. Drawing upon multivariate analyses, the results demonstrated that early language and literacy experiences as a whole contributed to later HL reading development. More specifically, language and literacy experiences in the community had a more salient effect on CHL literacy development than did home-based experiences. The study underscored the importance of crossing boundaries between learner-internal linguistic capacities and learner-external sociocultural contexts in HL literacy maintenance.

Keywords Oral and print input · Heritage language development and maintenance · Heritage language schools · Bilingual and biliteracy

H. Zhang (✉) · X. Cheng · J. Lin
The Psycholinguistics Lab, School of Foreign Languages, East China Normal University,
Shanghai, China
e-mail: hmzhang@english.ecnu.edu.cn; 51190400046@stu.ecnu.edu.cn;
52190400008@stu.ecnu.edu.cn

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics
58, https://doi.org/10.1007/978-3-031-24078-2_5

1 Introduction

1.1 The Tuckerian Impact in Research on Bilingual and Biliteracy Development

During the past decades, globalization has changed the bilingual or multilingual landscape at an ever-quickening pace. Bilingualism or multilingualism permeates everyday practices as individuals commonly use two or more languages (Dutcher & Tucker, 1996). Bilingual education is accorded much attention with an upsurge of research investigating its effectiveness. Prior to the 1980s, amid a time of the English-only movement, bilingual education was met with objections asserting that bilingual children would suffer mental retardation, lag behind their monolingually instructed peers, and sink into alienation (Tucker, 1989). Yet, Tucker, whom the present volume honors, made a clear articulation about bilingual advantages through innovative and dynamic instruction. As outlined in Tucker's research, prevailing beliefs about bilingual handicap were nothing but myths and misinterpretations of bilingualism (Tucker & d'Anglejan, 1971). Opposing a divisive and fragmented society where English dominated, he has provided empirical evidence for the cognitive and social advantages of foreign language education.

In the 1970s, Tucker, together with his colleagues, conducted an array of large-scale research projects exploring language teaching and language learning through the bilingual lens. One of his exemplary achievements was the longitudinal collaborative programs of the St. Lambert Experiment, in which multiple studies were conducted to examine the linguistic and academic progress of children from kindergarten through the elementary school years. A series of annual reports cumulatively demonstrated that bilingual immersion education constituted no impediment to children's academic, linguistic, or cognitive progress (see Bruck et al., 1974, 1977; Lambert & Tucker, 1972; Lambert et al., 1973; Tucker, 1975). Instead, the pupils in the experimental group were equally adept at reading, writing, speaking and using English with their counterparts in the control group. Furthermore, they were competent at learning French as a second language. Most striking was the considerable progress in social and cognitive development. Pupils educated in bilingual immersion programs developed better communication skills (Genesee et al., 1975). In addition, they exhibited high motivation and positive attitudes towards second language learning (Lambert et al., 1973). In this regard, Tucker's collaborative projects broke the long-held myths against bilingual education. These landmark studies lend empirical support to the effectiveness of language education programs, and attest to the benefits of bilingual and biliteracy development throughout the elementary school years.

Tucker is a pioneer of crossing the boundaries between bilingual education research and pedagogy. As a scholar practitioner, he provided insights into bilingual teaching grounded in theoretical observations and implementations. As noted in his review of a global perspective on bilingualism and bilingual education, the world is witnessing more bilingual than monolingual individuals, and education through a

second language is becoming more of a common scenario (Tucker, 1999). In a systematic review, he elaborated on the innovative approaches to bilingual schooling and stated that language education programs could bring advantages instead of cognitive and social detriments to children (Tucker, 1989). In addition to defining the essential criteria for exemplary bilingual immersion programs (such as instructional demands and language curriculums), his work shed light on diverse approaches to second language teaching (e.g., Tucker, 1974, 1977). Tucker exemplified bilingual immersion programs as a powerful vehicle to enhance bilingual competency throughout the early elementary school years. Tucker has devoted himself to crossing linguistic boundaries between being a researcher and a practitioner. He emphasized innovative immersion programs and developed research paradigms and pedagogical approaches in the field of bilingual and biliteracy education.

To highlight, in second language and foreign language learning, Tucker (1999) argues that linguistic resources, once developed and acquired, can be shared across languages. What factors may affect second language learning has been a thread running through his scholarship. In a broad array of publications, he examined learner-internal factors in second language learning, such as affective and cognitive variables, as well as learner-external factors at the contextual level (d'Angelian & Tucker, 1973; Lambert et al., 1973; Tucker et al., 1976). Tucker's research made two particularly salient notes on developing second language skills, that is, language input and language instruction. For instance, in a study that investigated Arabic-speaking English language learners, Tucker and colleagues discovered that the level of English language proficiency was contingent upon whether English was used as the medium of instruction (Saegert et al., 1974). This finding was further substantiated in Tucker's later research where systematic observations were conducted among pupils learning Japanese as a foreign language in the United States. The results revealed that over the seven-month span examined, classroom instruction affected children's later performance on linguistic proficiency measures (Donato et al., 2000). Furthermore, Tucker and his colleagues placed great emphasis on dual language input. In an examination of classroom input and teaching behaviors at school, Hamayan and Tucker (1980) found that certain forms of language production were affected by the frequency of occurrence of linguistic input and teaching strategies in the classroom setting. Tucker's research converges to suggest that an environment enriched with linguistic resources (i.e., an external factor) enhances second language development. In summary, Tucker envisioned that bilingual learning is a multifaceted construct shaped and aggregated by learners themselves and the surrounding environment.

The present chapter draws upon Tucker's vision of bilingualism and biliteracy, especially his examination of learner-internal and learner-external factors in bilingual learning. As Tucker (1999) expounded in his review, teacher instruction in language immersion programs, aided by parental involvement and community support, can pave the way for successful second language learning. Despite the foundation Tucker and colleagues have laid for understanding bilingual/biliteracy development in light of both internal and external factors, little research in language education has studied learner-external, input-based experiences in HL literacy

development. HL development not only depends on many learner-internal factors but also necessitates external support, including support from home and communities.

1.2 Home and Community Support in HL Maintenance

According to Valdés (2000), heritage language learners in the US are those who are raised in homes where a language other than English is used. Most heritage speakers have a good command of the majority language, but they face the challenge of developing HL literacy because of restricted print language input (Koda et al., 2008b). This may lead to weak minority language proficiency and bring difficulties for HL maintenance. Rather than being developed within the conventional K-12 educational system, HL literacy is often acquired through family education or community-based ethnic language schools (Kim & Pyun, 2014). Prior studies on bilingual and biliteracy development have demonstrated the importance of early language exposure in later HL language development (e.g., De Houwer, 2007; Schwartz, 2008; Thordardottir, 2011; H. Zhang, 2016; H. Zhang & Koda, 2018a), and HL input quantity (e.g., Daskalaki et al., 2018) and quality (e.g., y Cabo, 2020) are shown to be significantly related to success in HL acquisition. However, what remains unknown is the relative contributions of home and community support to HL literacy.

A number of studies have affirmed the predictive role of both home and community support in HL development and claimed that the support from home and community provided young HL learners with linguistic and metalinguistic foundations (Sun et al., 2016, 2020; H. Zhang, 2016). For example, H. Zhang (2016) investigated the relationship between early language input and later reading development among Chinese heritage language (CHL) learners. They were administered a background questionnaire and measurements of print vocabulary knowledge, lexical inferencing, and reading comprehension. Drawing upon regression analyses, the study found that oral and print language experiences at home and in CHL schools were related to HL reading skills. Likewise, Sun et al. (2020) investigated the effect of family factors and preschool factors on HL vocabulary knowledge development among 457 Singaporean preschool children whose HL was Mandarin, Malay, or Tamil. Their findings revealed that input quantity and quality at home and school predicted HL vocabulary growth.

Recent studies have identified various degrees of contributions of home and school environments to HL literacy development. Some highlighted the overwhelming effect of language input at school on heritage language development (e.g., Sun et al., 2016; Thordardottir, 2011), especially among the older heritage learner group. HL schools provide an ideal place for HL learners, where they can receive HL language instruction and cultivate ethnic cultural awareness through the construction of ethnic identity (Shibata, 2000). HL proficiency has been demonstrated to have positive connections with learners' involvement in the ethnic community (e.g.,

H. Zhang & Koda, 2018a). Sun et al. (2016) investigated the role of internal factors (age of English onset, short-term memory) and external factors including the input from schools and homes in English vocabulary development among 43 pre-school EFL learners in China. The results exhibited that the external factors played a crucial role in bilingual vocabulary development at an early stage, and the amount of English input at school facilitated the development of English syntactic knowledge. However, they found that the weekly English input at home did not exert a significant influence on any aspects of English vocabulary. Similarly, Druten-Frietman et al. (2015) investigated the relation between family and school factors in vocabulary growth among 385 preschoolers from native and non-native backgrounds in the Netherlands. The structural equation model revealed that family factors were crucial to early vocabulary development whereas school factors significantly predicted later vocabulary growth (Gerde & Powell, 2009).

There also have been a number of studies endorsing the positive effect of home input on HL language and literacy development (D. Zhang & Koda, 2011; H. Zhang & Koda, 2018a, 2018b). H. Zhang and Koda (2018a) highlighted the role of early oral exposure to home language in word-knowledge development by comparing CHL learners and non-CHL learners. CHL learners outperformed their non-CHL counterparts on oral comprehension ability (including oral vocabulary), lexical inference, and morphological awareness. Morphological awareness developed through oral language experiences may facilitate literacy development since it incorporates “binding agents” such as orthographic, phonological, grammatical, and semantic features that help to refine lexical representations and enhance lexical quality (Bowers et al., 2010, p.168). H. Zhang and Koda (2018a) confirmed that home language input could contribute to oral vocabulary and associate metalinguistic awareness with word-level skills. To be more precise, the study explored the interconnections between different word-level subskills among CHL learners. Oral language exposure had a significant effect on vocabulary knowledge, and oral vocabulary provided a solid linguistic foundation in facilitating morphological awareness, print vocabulary knowledge, and lexical inferencing ability. Koda et al. (2008a), in particular, underscored crossing boundaries between learner-internal, resource-related factors and learner-external, contextual factors in HL literacy development and maintenance.

Taken together, early home and community support is crucial to HL literacy development. However, a question may be raised about the extent language to which input, and what kind of language input, could support HL literacy skills. Koda et al. (2008a) identified print and oral proficiency as two important factors affecting CHL literacy development. Subsequently, H. Zhang and Koda (2018b) explored the diversities of early language input (print and oral-related activities) and their impact on word-knowledge development among adult CHL learners. They found that print-related activities were the dominant factor predicting CHL learners’ early language experiences. Those with more exposure to print at home tended to perform better in reading-related activities. And the intensity of the print experiences at school was the dominant factor distinguishing the low-input and high-input groups. The authors explained that print experiences may help learners establish the connections between

sound, graphic form, and meaning. The significant predictive power of reading practices at home was also verified in their prior studies (e.g., D. Zhang & Koda, 2011). Park et al. (2012) further explored the relative contributions of these factors by examining the transactional relationships between HL proficiency and family climate among Chinese pre-school children from immigrant families. Family climate was assessed in terms of parental warmth, cultural maintenance values, and use of HL support. The results revealed that the parental behavioral support of HL was closely associated with HL development. Concurrently, earlier HL proficiency was a significant predictor affecting subsequent parental behavior and those children with limited HL proficiency may affect their parents' use of HL support later.

In summary, caregivers and school factors play a significant role in HL development, and oral and print input in different social contexts may contribute to variations in HL language and literacy performance. However, to date, there is still a dearth of empirical research scrutinizing the impact of early language experiences on later HL literacy development, especially among CHL learners with diverse educational and social backgrounds. In addition, CHL learners develop different aspects (e.g., formality, register, and genre) of language and literacy skills in various social situations and it is worth examining the extent to which different types of input can affect HL literacy acquisition.

2 The Present Study

Early work in bilingual contexts has shown that bilingual linguistic resources can be shared across languages to enhance bilingual and biliteracy development (Cohen & Swain, 1976; Genesee et al., 1975; Tucker, 1977). The heritage language population creates a unique case to explore the role of linguistic resources in maintaining HL language and literacy skills. HL learners vary widely in their early language and literacy experiences and, as a result, they develop different oral and written communication skills. Prior language and literacy support from home, the community, and HL schools provides young HL learners with linguistic and metalinguistic foundations for language and literacy development (Chinen & Tucker, 2005; Koda et al., 2008a; Mori & Calder, 2013). Given that the Tuckerian impact in bilingual/biliteracy development underscores both internal and external factors, the current study aims to explore the role of learner-external and input-based language and literacy support in adult CHL literacy development. The overarching research question addressed in the current study is how early language and literacy experiences at home and in the community contribute to later CHL literacy development. More specifically, this study intends to compare home language input and community language support in CHL literacy development, thus highlighting the respective contributions of different types of input to CHL literacy development.

3 Method

3.1 *Participants*

Participants were from a dataset collected from 2015 to 2018. A total of 209 adult CHL students (mean age = 20.32 years, 139 females and 70 males) participated in this large-scale study. During the data collection, the participants were from English-speaking countries and they were all enrolled in study-abroad programs in Beijing, Shanghai, and Guangzhou, China. All the participants had at least one year of intensive classroom learning experience. Their instruction levels ranged from low intermediate to high intermediate. They had extended exposure to Chinese, including Mandarin and regional varieties during their childhood.

3.2 *Instruments*

3.2.1 Home and Community Language and Literacy Experiences

A retrospective questionnaire (presented in English) was used to elicit CHL students' responses to early language exposure at home and in the community. The questionnaire was modelled after Shu et al. (2002) and Xiao (2008), which emphasized oral language and print literacy experiences in a variety of sociocultural contexts (home, ethnic community, and after-school/weekend HL school). The questions asked about the frequency and intensity of language and literacy experiences in various social situations.

Oral Language Experiences

There were 27 questions in the question pool regarding oral language experiences. Questions about contextual support included oral language use at home ($k^1 = 8$) and in the community ($k = 8$). The community, more specifically, included both the ethnic community and HL schools. For example, the participants were asked what language they used to communicate in their family and in the ethnic community. Each item was normed based on a 1-to-5 Likert scale (1 for English only, 2 for mostly English, 3 for half Chinese half English, 4 for mostly Chinese, 5 for Chinese only).

Print Literacy Experiences

Print literacy experiences focused on the three parameters of Shu et al. (2002): literacy resources at home, parent-child literacy activities and independent activities. Likewise, questions were asked about print literacy experiences at home ($k = 7$) and in the community (mostly in HL schools) ($k = 3$).

¹ k : number of items.

3.2.2 Reading Vocabulary Knowledge

The vocabulary measurement asked the participants to identify appropriate semantic meanings of visually-presented words. A total of 30 multiple-choice questions were shown to the participants and they needed to choose the accurate meaning of each presented Chinese character or word. For example, the word 抱怨 (complain) along with 4 options (1. hug, 2. complain, 3. discuss, and 4. explain) was presented and the participants were required to choose the accurate meaning from the four choices. The reliability was $\alpha = .818$.

3.2.3 Morphological Awareness

The morphological awareness measures tapped into students' graphomorphological knowledge, which was operationalized as morpheme segmentation and morpheme discrimination (40 items in total). The segmentation task required the participants to break down bimorphemic three-character words into two parts. For example, 出租车 (literally out-rent-car; taxi) can be broken down into 出租 (rent out) and 车 (vehicle). The discrimination task, which was adapted from Ku and Anderson (2003), aimed to measure the participants' ability to analyze and discriminate morphemic structures and meanings. For example, three words were presented: 读者 (reader), 学者 (scholar), 或者 (or). The shared character “者” refers to “person or professional” in most cases, however 或者 (or) does not share the same morphemic meaning with the other two. The participants were asked to circle the odd one out. The reliability was $\alpha = .727$.

3.2.4 Lexical Inferencing Ability

The lexical inferencing ability task tested the participants' ability to infer meanings of unfamiliar words based on word-internal (morphological/character knowledge) and word-external (contextual cues) information. Disyllabic compound words were used as stimuli in the measure. Each word consisted of two elementary-level Chinese characters known to the participants. However, the combination was unknown to the intermediate level learners given the frequency and familiarity were beyond the highest level of the HSK, a standardized proficiency test for non-native speakers of Chinese (see Peng et al., 2021). Before the actual testing, 14 level-appropriate Chinese learners completed a pilot test to rate the familiarity of 20 initially selected compound words. According to the results of the pilot, we finalized 16 disyllabic words as target words in the actual test. For example, the short sentence 他是一个钢琴高手 (He is a piano _____) was presented and four options for the meaning of 高手 (1. expert, 2. teacher, 3. high-handed, 4. left-handed) were provided. The participants were supposed to utilize both word-internal and word-external cues to derive the meaning (expert). The reliability was $\alpha = .856$.

3.2.5 Reading Comprehension

The measure for reading comprehension was adopted from a HSK reading test. Based on the proficiency level of the participants, intermediate level (HSK 3 to 5) comprehension questions were administered to the participants. Questions were presented with a mixture of questions from different levels. Each short passage was followed by one multiple-choice comprehension question. Question types included co-reference building, gist detection, text meaning inference, and specific information identification. There was a total of 18 passages along with 18 questions. The reliability was $\alpha = .763$.

4 Results

4.1 Descriptive Statistics

Table 1 presents the descriptive results of the tested variables. The results showed that the participating CHL students had wide variability with regard to oral and print experiences because the indices of oral language and print literacy experiences in various sociocultural contexts had relatively wide spread, which suggested that the learners had different degrees of early language exposure. The accuracy rates of CHL literacy measurements ranged from 75.4% (vocabulary knowledge) to 83.8% (morphological awareness). The measurements of CHL literacy skills also had relatively wide dispersions given the standard deviations, which indicated that the CHL students had heterogeneous linguistic and literacy profiles.

Table 1 Descriptive statistics of oral and print experiences and CHL literacy

Variable	M	SD	Min	Max
1. Oral language input at home (40)	15.75	7.48	8	40
2. Oral language input in the community (40)	14.84	6.41	8	40
3. Print literacy input at home (35)	13.29	5.24	7	32
4. Print literacy input in the community (15)	6.87	4.23	3	15
5. CHL vocabulary knowledge (30)	22.63	5.00	5	30
6. CHL morphological awareness (40)	33.52	4.68	12	40
7. CHL lexical inference (16)	12.06	2.75	3	16
8. CHL reading comprehension (18)	14.29	3.48	4	18

Note. Numbers in parentheses indicate the maximum for each variable

4.2 Bivariate Correlations

Table 2 presents the results of bivariate correlations between early language experiences and later CHL literacy skills. Within the indicators of early language and literacy experiences, four variables were significantly correlated with each other. The correlations ranged from $r = .241, p < .01$ to $r = .580, p < .001$. The correlations were overall moderate in size between the experience-related variables within an input modality (oral versus print) or a context (home versus community). More important, Table 2 shows that oral language experiences at home and in the community had significant correlations with CHL vocabulary, morphological awareness, and lexical inference ($r = .173, p < .05$ to $r = .267, p < .01$). Print literacy experiences at home also had significant correlations with CHL vocabulary and lexical inference ($r = .147, p < .05$ and $r = .190, p < .01$, respectively), and print literacy experiences in the community had significant correlations with all four literacy measurements ($r = .189, p < .05$ to $r = .292, p < .01$). It is worth noting that CHL reading comprehension was only correlated with print literacy experiences in the community.

4.3 Multiple Regressions

Multiple regression was employed to test the individual and collective contributions of early language input at home and in the community to CHL literacy development. Oral language indicators and print literacy indicators were entered into the

Table 2 Correlational matrix of oral and print experiences and CHL literacy

	1	2	3	4	5	6	7	8
1. Oral language input at home (40)	–							
2. Oral language input in the community (40)	.492***	–						
3. Print literacy input at home (35)	.419***	.347***	–					
4. Print literacy input in the community (15)	.241**	.401**	.580***	–				
5. CHL vocabulary knowledge (30)	.195*	.175*	.147*	.236**	–			
6. CHL morphological awareness (40)	.173*	.259**	.128	.189**	.651***	–		
7. CHL lexical inference (16)	.204**	.267**	.190**	.292***	.603***	.590***	–	
8. CHL reading comprehension (18)	.090	.132	.090	.215**	.599***	.632***	.636***	–

* $p < .05$; ** $p < .01$, *** $p < .001$

two blocks of regression models (Table 3). In Model 1, both oral language and print literacy input contributed to CHL reading vocabulary knowledge. Oral language input predicted 4.6% of the variance in vocabulary knowledge.

Strikingly, an additional 3.2% of the variance was predicted by print literacy input after the effect of oral language input was incorporated. Model 2 demonstrated that oral language input predicted CHL morphological awareness (7.0% of the total variance explained) whereas print literacy input did not have a significant impact on CHL morphological awareness after oral language input was accounted for. Model 3 showed that oral language input and print literacy input both contributed to CHL lexical inference. Oral language input predicted 7.8% of the variance in lexical inference and print literacy input uniquely predicted 3.9% of the variance after oral language input was taken into account. The last model predicted that print literacy input had a significant effect on CHL reading comprehension given that print literacy input predicted a unique proportion of variance in reading comprehension (3.5%

Table 3 Hierarchical regressions predicting CHL literacy skills

	Model 1: CHL vocabulary knowledge		Model 2: CHL morphological awareness		Model 3: CHL lexical inference		Model 4: CHL reading comprehension	
	β	t	β	t	β	t	β	t
<i>Chinese oral language input</i>	.14	1.78	.06	.75	.10	1.21	.03	.39
Oral input at home	.11	1.29	.23	2.87**	.22	2.76**	.12	1.42
Oral input in the community								
<i>Chinese print literacy input</i>								
Print input at home	-.05	-.58	-.03	-.35	-.03	-.35	-.08	-.85
Print input in the community	.22	2.45*	.11	1.28	.23	2.66**	.23	2.67*
R^2	.046 (.079)		.070 (.078)		.078(.118)		.018 (.053)	
ΔR^2	.046 (.032)		.070 (.008)		.078 (.039)		.018 (.035)	
SE	4.67 (4.62)		4.30 (4.30)		2.63 (2.58)		3.42 (3.37)	
ΔF	4.68* (3.32*)		7.21** (.89)		8.15*** (4.25*)		1.79 (3.47*)	

* $p < .05$; ** $p < .01$; *** $p < .001$. Note. Numbers in the parentheses refer to the statistical indices of print experience

of the variance). No significant effect of oral language input was found in reading comprehension even when it was entered first.

Probing into the individual contributions of input experiences at home and in the community, we can find, based on the regression coefficients presented in Table 3, that language experiences in the community outweighed language experiences at home in predicting CHL literacy skills. In the final regression models where all input experience indicators were included, the significant unique predictor(s) was always the experience in the community. Specifically, oral language input in the community predicted CHL morphological awareness and lexical inference ($t = 2.87$, $p < .01$ and $t = 2.76$, $p < .01$, respectively). Furthermore, print literacy input contributed to CHL vocabulary, lexical inference, and reading comprehension ($t = 2.45$, $p < .05$; $t = 2.66$, $p < .01$; and $t = 2.67$, $p < .05$, respectively).

5 Discussion

5.1 Input Modality in CHL Literacy Development

The study underscored both oral language input and print literacy input in later CHL literacy development. The results demonstrated that oral language experiences had significant correlations with vocabulary, morphological awareness, and lexical inference but not reading comprehension. Oral language input, in general, contributed to the development of word-level literacy skills. H. Zhang and Koda (2018a) found that early oral language exposure enhanced word-level skills and morphological awareness, which was consistent with the findings of the current study. However, early oral language input alone does not seem to have a long-term effect on higher-level reading comprehension because discourse comprehension not only involves decoding abilities but global text-based meaning construction. It is also worth noting that oral language input outweighed print literacy input in predicting Chinese morphological awareness. The current study corroborated that morphological awareness was more strongly influenced by oral language experiences among CHL students.

Earlier studies have shown that morphological knowledge and awareness start to develop once children have oral communication (Berko, 1958), and acquisition of fundamental morphological structures is achieved by early elementary age (Anisfeld & Tucker, 1968; Berko, 1958). Morphological awareness is a multilayered construct in literacy acquisition that encompasses phonology, orthography, and semantics (Kuo & Anderson, 2006). Abstraction of morphemic meanings and morphological structures can be found in both phonological and graphic representations. The literature on monolingual children has found that oral language competence is closely tied to morphological awareness. McBride-Chang et al. (2005) explored the relationship between morphological awareness and oral vocabulary knowledge in L1 English-speaking children. Morphological structure awareness and morpheme

identification, as two indicators of morphological awareness, were found to have significant correlations with oral vocabulary knowledge. More strikingly, McBride-Chang et al. (2008) highlighted the bidirectional relationship between morphological awareness and oral vocabulary in Mandarin, Cantonese, and Korean. Oral language contributed to morphological awareness, and emerging morphological awareness affected oral vocabulary.

In the CHL population, Koda et al. (2008a) found an effect of print input on Chinese morphological awareness in CHL students, given that radical segmentation and analysis in textbook vocabulary can enhance morphological decomposition. Although the current study tapped into graphomorphological awareness, it was notable that the connection of this morphological awareness with oral language experience was more salient. The current study indicates that early oral language experiences can have a lasting and unique impact on morphological awareness in CHL students. Morphological awareness developed through oral language experiences may foster literacy development given that it integrates phonological, orthographic, grammatical, semantic features to refine lexical representations and quality (Bowers et al., 2010).

Oral and print resources, as external factors, enhance learner-internal reading subskills including morphological awareness, reading vocabulary, lexical inference and reading comprehension. Although oral language experiences do not directly enhance higher-level reading comprehension, they shape the formation of fundamental metalinguistic awareness, which further fosters higher-level reading abilities.

5.2 Community Language and Literacy Support in CHL Literacy

Literacy is not just the linguistic ability to read and write. More important, literacy learning is situated in broader social contexts. The concept of plural literacies connects learner-internal factors with learners' experiences with the world (Freire & Macedo, 1987) and literacies integrate both individual and social events (Behrman, 2002; Haneda, 2006). The current study highlighted environmental factors in light of the effect of linguistic input in different social contexts, that is, home and community, on later CHL literacy development. Community language and literacy support, in particular, was found to be associated with all the measured CHL literacy skills. It is important to discuss the role of community support in CHL literacy development and maintenance.

First of all, communities establish a dynamic context to engage learners in collaborative and interactive learning (Gregory, 2001; Williams & Gregory, 2001; Zhou & Kim, 2006). Literacy acquisition necessitates the interaction between learners and communities. Communities can be divided into three different categories: *classroom community*, *experiential community* and *anticipatory community* (Behrman, 2002). The latter two communities refer to learning in real-life contexts,

in which literacy learners and communities build an active interaction. Gregory (2001) explored heritage language and literacy development among multilingual Bangladeshi families in East London. In conjunction with the existing findings, the current study suggests that home and community learning environments create a synergy of reciprocal and collaborative learning.

Second, ethnic community and after-school HL programs help to construct HL learners' ethnic identity as well as cultural and linguistic capital (Chinen & Tucker, 2005; Pu, 2010; Shibata, 2000). Ethnic community and after-school programs foster the (re)connection with ethnic identity. Chinen and Tucker (2005) investigated learners of Japanese as a heritage language in a weekend HL school in Los Angeles. Over the course of six months, they found that HL children had stronger cultural bonds and heritage identity. More important, their positive attitudes toward heritage culture and identity were related to their language and literacy performance. Pu (2010) also found that CHL children developed distinct understandings of cultural and linguistic capital in HL schools, which enhanced their literacy learning in a contextualized environment.

Third, multiple communities of practice are formed in HL schools to expand learners' linguistic repertoires. After-school literacy programs can enable teachers, students, parents, and school personnel to work collaboratively, and literacy-related activities provide interaction between communities of practice that are different from home and school (Peercy et al., 2013; Sneddon, 2000). Limited exposure to various genres or registers at home may not be conducive to integrated literacy development. Sneddon (2000) found that oral language at home did not have a significant effect on biliteracy development among HL learners. This was similarly the case in the present study (see Table 3). Given the ethnolinguistic vitality, active participation in various communities of practice in the ethnic community or HL schools can facilitate literacy performance. Our questionnaire asked the participants to identify the frequency and intensity of Chinese literacy-related activities, for example, practicing calligraphy or reading rhyming poems. The results demonstrated that early literacy-related activities in HL schools were significantly correlated with later CHL literacy. Literacy-related activities in HL schools can engage students in a socioculturally-rich environment, thus expanding their cultural and linguistic repertoire. Early literacy exposure in HL schools can shape and influence later HL literacy development.

6 Conclusion

In this chapter, we reported on our empirical investigation of early language input at home and in the community as well as the relationship between early input and later HL literacy development. Using a retrospective questionnaire and multiple literacy measurements, the study found that early input experiences in general had positive effects on later HL literacy development. First, input modality affected the strength of correlations between early experiences and later HL literacy skills. More

specifically, morphological awareness was only affected by oral language input while the unique proportion of variance in reading comprehension ability was solely predicted by print literacy input. Second, the study showed that compared with home input experiences, community language and literacy support had a more salient effect on HL literacy development. The ethnic community and HL schools can build the bond between learners and cultural heritage, and also create communities of practice that are different from home and school, which can enhance students' positive and active participation in literacy practices. An environment enriched with external linguistic resources can enhance bilingual and biliteracy development (d'Angelian & Tucker, 1973; Lambert et al., 1973; Tucker et al., 1976). The study underscored the importance of crossing boundaries between learner-internal linguistic capacities and learner-external sociocultural contexts in HL literacy maintenance.

References

- Anisfeld, M., & Tucker, G. (1968). English pluralization rules of six-year-old children. *Child Development, 38*, 1201–1217.
- Behrman, E. H. (2002). Community-based literacy learning. *Literacy, 36*(1), 26–32.
- Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150–177.
- Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*, 144–179.
- Bruck, M., Lambert, W. E., & Tucker, G. R. (1974). Bilingual schooling through the elementary grades: The St. Lambert project at grade 7. *Language Learning, 24*(2), 183–204.
- Bruck, M., Lambert, W. E., & Tucker, G. R. (1977). Cognitive consequences of bilingual schooling: The St. Lambert project through grade six. *Linguistics, 15*, 13–33.
- Chinen, K., & Tucker, G. R. (2005). Heritage language development: Understanding the roles of ethnic identity and Saturday school participation. *Heritage Language Journal, 3*(1), 27–59.
- Cohen, A. D., & Swain, M. (1976). Bilingual education: The “immersion” model in the north American context. *TESOL Quarterly, 10*, 45–53.
- d'Angelian, A., & Tucker, G. R. (1973). Communicating across cultures: An empirical investigation. *Journal of Cross-Cultural Psychology, 4*(1), 121–130.
- Daskalaki, E., Chondrogianni, V., Blom, E., Argyri, F., & Paradis, J. (2018). Input effects across domains: The case of Greek subjects in child heritage language. *Second Language Research, 36*(5), 421–445.
- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied PsychoLinguistics, 28*(3), 411–424.
- Donato, R., Tucker, G. R., Wudthayagorn, J., & Igarashi, K. (2000). Converging evidence: Attitudes, achievements, and instruction in the later years of FLES. *Foreign Language Annals, 33*(4), 377–393.
- Druten-Frietman, L. V., Denessen, E., Gijssels, M., & Verhoeven, L. (2015). Child, home and institutional predictors of preschool vocabulary growth. *Learning and Individual Differences, 43*, 92–99.
- Dutcher, N., & Tucker, G. R. (1996). *The use of first and second languages in education: A review of international experience*. Retrieved from <https://documents1.worldbank.org/curated/en/131161468770987263/pdf/multi-page.pdf>
- Freire, P., & Macedo, D. (1987). *Literacy: Reading the word and the world*. Bergin and Garvey.

- Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communication skills of bilingual children. *Child Development*, 46(4), 1010–1014.
- Gerde, H. K., & Powell, D. R. (2009). Teacher education, book-reading practices, and children's language growth across one year of head start. *Early Education and Development*, 20, 211–237.
- Gregory, E. (2001). Sisters and brothers as language and literacy teachers: Synergy between siblings playing and working together. *Journal of Early Childhood Literacy*, 1(3), 301–322.
- Hamayan, E., & Tucker, G. R. (1980). Language input in the bilingual classroom and its relationship to second language achievement. *TESOL Quarterly*, 14(4), 453–468.
- Haneda, M. (2006). Becoming literate in a second language: Connecting home, community, and school literacy practices. *Theory Into Practice*, 45(4), 337–345.
- Kim, C. E., & Pyun, D. O. (2014). Heritage language literacy maintenance: A study of Korean-American heritage learners. *Language, Culture and Curriculum*, 27(3), 294–315.
- Koda, K., Lü, C., & Zhang, Y. (2008a). Effects of print input on morphological awareness among Chinese heritage language learners. In A. W. He & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 125–135). University of Hawai'i, National Foreign Language Resource Center.
- Koda, K., Zhang, Y., & Yang, C.-L. (2008b). Literacy development in Chinese as a heritage language. In A. W. He & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 137–149). University of Hawai'i, National Foreign Language Resource Center.
- Ku, Y.-M., & Anderson, R. C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing: An Interdisciplinary Journal*, 16(5), 399–422.
- Kuo, L.-J., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*, 41(3), 161–180.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Newbury House Publishers.
- Lambert, W. E., Tucker, G. R., & d'Anglejan, A. (1973). Cognitive and attitudinal consequences of bilingual schooling. *Journal of Educational Psychology*, 65(2), 141–159.
- McBride-Chang, C., Wagner, R. K., Muse, A., Chow, B. W., & Shu, H. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied PsychoLinguistics*, 26(3), 415–435.
- McBride-Chang, C., Tardif, T., Cho, J. R., Shu, H., Fletcher, P., Stokes, S. F., & Leung, K. (2008). What's in a word? Morphological awareness and vocabulary knowledge in three languages. *Applied PsychoLinguistics*, 29(3), 437–462.
- Mori, Y., & Calder, T. M. (2013). Bilingual vocabulary knowledge and arrival age among Japanese heritage language students at hoshuukoo. *Foreign Language Annals*, 46(2), 290–310.
- Park, H., Tsai, K. M., Liu, L. L., & Lau, A. S. (2012). Transactional associations between supportive family climate and young children's heritage language proficiency in immigrant families. *International Journal of Behavioral Development*, 36(3), 226–236.
- Peercy, M. M., Martin-Beltran, M., & Daniel, S. M. (2013). Learning together: Creating a community of practice to support English language learner literacy. *Language, Culture and Curriculum*, 26(3), 284–299.
- Peng, Y., Yan, W., & Cheng, L. (2021). Hanyu Shuiping Kaoshi (HSK): A multi-level, multi-purpose proficiency test. *Language Testing*, 38(2), 326–337.
- Pu, C. (2010). The influence of heritage language and public schools on Chinese American children's biliteracy development. *Bilingual Research Journal*, 33(2), 150–172.
- Saegert, J., Scott, S., Perkins, J., & Tucker, G. R. (1974). A note on the relationship between English proficiency, years of language study and medium of instruction. *Language Learning*, 24(1), 99–104.
- Schwartz, M. (2008). Exploring the relationship between family language policy and heritage language knowledge among second generation Russian-Jewish immigrants in Israel. *Journal of Multilingual and Multicultural Development*, 29, 400–418.

- Shibata, S. (2000). Opening a Japanese Saturday school in a small town in the United States: Community collaboration to teach Japanese as a heritage language. *Bilingual Research Journal*, 24(4), 465–474.
- Shu, H., Li, W., Anderson, R., Ku, Y.-M., & Yue, X. (2002). The role of home-literacy environment in learning to read Chinese. In W. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition: Theoretical and pedagogical issues* (pp. 207–224). Kluwer Academic Publishers.
- Sneddon, R. (2000). Language and literacy: Children's experiences in multilingual environments. *International Journal of Bilingual Education and Bilingualism*, 3(4), 265–282.
- Sun, H., Steinkrauss, R., Tendeiro, J., & de Bot, K. (2016). Individual differences in very young children's English acquisition in China: Internal and external factors. *Bilingualism: Language and Cognition*, 19(3), 550–566.
- Sun, H., Ng, S. C., O'Brien, B. A., & Fritzsche, T. (2020). Child, family, and school factors in bilingual preschoolers' vocabulary development in heritage languages. *Journal of Child Language*, 47(4), 817–843.
- Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, 15, 426–445.
- Tucker, G. R. (1974). *Innovative approaches to second-language teaching*. TESOL.
- Tucker, G. R. (1975). The acquisition of knowledge by children educated bilingually. In D. P. Dato (Ed.), *Georgetown University round table on languages and linguistics 1975* (pp. 267–277). Georgetown University Press.
- Tucker, G. R. (1977). Can a second language be taught? In H. D. Brown, C. A. Yorico, & R. H. Cryme (Eds.), *Teaching and learning English as a second language: Trends in research and practice* (pp. 14–30). TESOL.
- Tucker, G. R. (1989). Cognitive and social correlates of bilinguality. In R. Cooper & B. Spolsky (Eds.), *The influence of language on culture and thought: Essays in honor of Joshua A. Fishman's sixty-fifth birthday* (pp. 1016–1111). Mouton de Gruyter.
- Tucker, G. R. (1999). *A global perspective on bilingualism and bilingual education*. ERIC Clearinghouse on Languages and Linguistics. Retrieved from <https://files.eric.ed.gov/fulltext/ED435168.pdf>
- Tucker, G. R., & d'Anglejan, A. (1971). Some thoughts concerning bilingual education programs. *Modern Language Journal*, 55(8), 4916–4493.
- Tucker, G. R., Hamayan, E., & Genesee, F. H. (1976). Affective, cognitive and social factors in second-language acquisition. *Canadian Modern Language Review*, 32(3), 214–226.
- Valdés, G. (2000). The teaching heritage languages: An introduction for Slavic-language-teaching professionals. In O. Kagan & B. Rifkin (Eds.), *The learning and teaching of Slavic languages and cultures* (pp. 375–403). Slavica.
- Williams, A., & Gregory, E. (2001). Siblings bridging literacies in multilingual contexts. *Journal of Research in Reading*, 24(3), 248–265.
- Xiao, Y. (2008). Home literacy environment in Chinese as a heritage language (CHL) development. In A. W. He & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 151–166). University of Hawai'i, National Foreign Language Resource Center.
- Y Cabo, D. P. (2020). Examining the role of cross-generational attrition in the development of Spanish as a heritage language: Evidence from gustar-like verbs. *Linguistic Approaches to Bilingualism*, 10(1), 86–108.
- Zhang, H. (2016). Early language input and later reading development in Chinese as heritage language (CHL) learners. *International Journal of Applied Linguistics*, 26(3), 437–448.
- Zhang, D., & Koda, K. (2011). Home literacy environment and word knowledge development: A study of young learners of Chinese as a heritage language. *Bilingual Research Journal*, 34, 4–18.
- Zhang, H., & Koda, K. (2018a). Word-knowledge development in Chinese as a heritage language learners: A comparative study. *Studies in Second Language Acquisition*, 40(1), 201–223.

- Zhang, H., & Koda, K. (2018b). Heterogeneity of early language experiences and word-knowledge development in Chinese as heritage language learners. In X. Wen & X. Jiang (Eds.), *Studies on learning and teaching Chinese as a second language* (pp. 14–37). Routledge.
- Zhou, M., & Kim, S. (2006). Community forces, social capital, and educational achievement: The case of supplementary education in the Chinese and Korean immigrant communities. *Harvard Educational Review*, 76(1), 1–29.

Part II

Teachers and Instructional Processes

Introduction and the Tuckerian Impact

This part is broadly defined to focus on any issues concerning the teacher, teaching and instructional process, as well as teacher education and teacher learning. Teachers are inarguably a key player in language education: what they know and think and what they do have critical implications on student learning (Part I) and the success of any language program (Part III), and they are a key policy actor in language education planning (Part IV). Accordingly, it is very important to cross boundaries to research and understand the teacher (who they are, what they know, and how they think) and their professional life (teaching and professional development), and to innovate classroom teaching and teacher learning. Language education researchers are always strongly interested in teaching methods and instructional processes, including, for example, classroom discursive processes and language learning. There has also been increasing interest in language teaching (Long & Doughty, 2009), language teacher cognition (Borg, 2006; Mercer & Kostoulas, 2018), and language teacher professional learning (Crandall, 2000; Crandall & Christison, 2016).

In fact, traditional understandings of “language teachers,” “language classrooms,” and “language instruction” are also being constantly redefined. For example, with the increasing representation of non-native English-speaking students in US schools, all teachers, including subject teachers, are expected to accommodate the language-related needs of culturally and linguistically diverse students (de Oliveira & Yough, 2015). Likewise, language teachers in CLIL (Content and Language Integrated Learning) or language immersion programs need to bridge language and content learning (Lo, 2020). Additionally, with the advent of new technologies, new environments of language instruction are emerging, such as blended learning or flipped classrooms, which are also redefining teachers’ roles, instructional practices, and processes of language learning (Lan, 2020; Mehring & Leis, 2018).

Boundaries need to be crossed, between languages, instructional contexts, learning environments, learners, and teaching methodologies and approaches, to

understand many issues around the teacher and teaching. These issues are also an integral and important component of boundary crossing in Dick's scholarship. Dick has published prolifically on language teaching, teachers, and teacher training and education, and crossed boundaries for innovative teaching and teacher education, whether these publications were contextualized in program implementation and evaluation (e.g., Campbell et al., 1974; see also Part III), policy engagements (e.g., Tucker, 1993; see also Part IV), or examination of the effects on language learning (e.g., Hamayan & Tucker, 1980; see also Part I).

Not only did Dick engage in discussions and debates on teaching approaches and methods, including content-based instruction (e.g., Tucker, 1974; Tucker & Crandall, 1989), but he conducted a number of studies with collaborators and graduate students to probe into classroom / instructional processes and student learning, crossing boundaries between languages, research methods, programmatic contexts, and learners. Those studies included, for example, quantification of linguistic features in teacher input and teaching behaviors (e.g., different strategies on errors) and analysis of their influence on student output in French immersion classrooms in Canada (Hamayan & Tucker, 1980), qualitative analysis of Spanish teachers' discursive practices and their implications on the learning of both language and content in content-based Spanish classrooms in US schools (Pessoa et al., 2007), and descriptions of the teacher's use of activities that gave students opportunities for creative writing in an elementary school Japanese-as-a-foreign-language classroom in the US (Mitsui et al., 2009).

In his commentary on the "big debate" on whether TESOL is an art or a science, which was entitled *Precision, elegance and simplicity: Perspectives on TESOL and Art*, Dick used the analogy of TESOL as a "Necker Cube" and convincingly argued that "in order to affect public policy, in order to improve the quality of education for the children and the adults who are entrusted to our care - that we must be both scientists and artists" (Tucker, 2000; p. 25), and we must cross boundaries to document "effective practice" as well as "the contributions of competing social, political, economic, and pedagogical factors on the processes and the products of teaching and learning;" aim for "accuracy, systematicity, and accessibility in our findings;" and "encourage work that is contextualized and that is collaborative—work that draws upon a broad range of models and analytic techniques" (Tucker, 2000; p. 26).

Another central thread of Dick's scholarship concerning the teacher is on teachers themselves and their learning or professional development. In a number of publications, Dick and his collaborators studied the psychology of teachers (e.g., self-concept, value orientations, professional motivations; e.g., Tucker & Lambert, 1970) and teachers' views on or attitudes toward diverse language education related issues, including notably policy and practice in the education of culturally and linguistically diverse students (e.g., Campbell et al., 1974; Tucker, 1993). They also studied the teacher knowledge base for language teaching (e.g., Acheson et al., 1978; Cowan et al., 1979), and compared methods or models for language teacher training and education (e.g., Bruck et al., 1975; Tucker & Lambert, 1970).

Collectively these studies crossed boundaries of countries and educational systems, such as China (Cowan et al., 1979), France (Tucker & Lambert, 1970), the Philippines (Bruck et al., 1975), Canada (Acheson et al., 1978; Campbell et al., 1974;), and the United States (Bruck et al., 1975; Tucker & Lambert, 1970). They used a variety of approaches and designs, such as questionnaire surveys on teachers' views and attitudes (Acheson et al., 1978; Campbell et al., 1974) and experimentation and assessment of teachers' psychological attributes (Bruck et al., 1975; Tucker & Lambert, 1970). Bruck et al. (1975) was perhaps one of the earliest studies that used longitudinal evidence to show the long-term benefits of studying/training abroad on teacher development and language teaching, despite the short-term evidence, such as participants'/trainees' immediate post-training dissatisfaction and depression, that painted an unfavorable picture about that training experience (see also Bowen, 1968). The findings strongly underpin the many more contemporary policies and practices in teacher education programs in the US and beyond that underscore study/experience abroad and global/multicultural education for teachers as well as internationalization of teacher education (e.g., AACTE, 1994; Shiveley & Misco, 2015; Walters et al., 2009).

Another salient and notable boundary crossing in Dick's scholarship on teachers and teacher education/learning is his envisioning of the teacher-researcher identity and relationship. Dick was among the earliest to make compelling arguments for the importance of a second language research component in language teacher education programs and outlined many ideas for providing research training to pre-service teachers and promoting their research engagement (Tucker, 1983; Tucker & Donato, 1995). He also actively collaborated with practitioners or in-service teachers and facilitated their crossing of boundaries as they became "researchers" of teaching innovation in their own classrooms (e.g., Mitsui et al., 2009) (see also Borg, 2010, 2013; Crandall, 2000). Dick's research has also provided insights into the complexity and fluidity of the researcher-practitioner (applied linguist - educator) identity, and the confusion it is associated with and the challenges it often brings, in the nexus of or the boundary between language education *research* and language *teaching* (Donato et al., 2014). (What was reported in Donato et al. [2014] about professional identities in applied linguistics, by the way, is perhaps what many if not all of the contributors in this volume have experienced, negotiated, and navigated under Dick's mentorship, as mentioned earlier in Chapter 1.)

References

- AACTE. (1994). *Teacher education in global & international education*. American Association of Colleges for Teacher Education.
- Acheson, P., d'Anglejan, A., de Bagheera, J., & Tucker, G. R. (1978). English as the Second Language in Quebec: A teacher profile. *McGill Journal of Education*, 13(2), 189–197.

- Borg, S. (2006). *Teacher cognition and language education: Research and practice*. Continuum.
- Borg, S. (2010). Language teacher research engagement. *Language Teaching*, 43(4), 391–429.
- Borg, S. (2013). *Teacher research in language teaching: A critical analysis*. Cambridge University Press.
- Bowen, J. D. (1968). *Final technical report on the Institute in English for Speakers of Other Languages*. University of California.
- Bruck, M., Lambert, W. E., Tucker, G. R., & Bowen, J. D. (1975). The 1968 NDEA Philippine institute for TESL teachers: A follow-up evaluation. *Foreign Language Annals*, 8(2), 133–137.
- Campbell, R. N., Taylor, D. M., & Tucker, G. R. (1974). Teachers' views of immersion-type bilingual programs: A Quebec example. *Foreign Language Annals*, 7(5), 106–110.
- Cowan, J. R., Light, R. L., Mathews, B. E., & Tucker, G. R. (1979). English teaching in China: A recent survey. *TESOL Quarterly*, 13(4), 465–482.
- Crandall, J. (2000). Language teacher education. *Annual Review of Applied Linguistics*, 20, 34–55.
- Crandall, J., & Christison, M. (Eds.). (2016). *Teacher education and professional development in TESOL: Global perspectives*. Routledge.
- de Oliveira, L. C., & Yough, M. (Eds.). (2015). *Preparing teachers to work with English Language Learners in mainstream classrooms*. Information Age.
- Donato, R., Tucker, G. R., & Hendry, H. (2014). Developing professional identities in applied linguistics: From doctoral study to professional practice. In Y. L. Cheung, S. B. Said, & K. Park (Eds.), *Advances and current trends in language teacher identity research* (pp. 217–234). Routledge.
- Hamayan, E., & Tucker, G. R. (1980). Language input in the bilingual classroom and its relationship to second language achievement. *TESOL Quarterly*, 14(4), 453–468.
- Lan, Y.-J. (2020). Immersion, interaction, and experience-oriented learning: Bringing virtual reality into FL learning. *Language Learning & Technology*, 24(1), 1–15.
- Lo, Y. Y. (2020). *Professional development of CLIL teachers*. Springer.
- Long, M. H., & Doughty, C. J. (Eds.). (2009). *The handbook of language teaching*. Wiley-Blackwell.
- Mehring, J., & Leis, A. (Eds.). (2018). *Innovations in flipping the language classroom*. Springer.
- Mercer, S., & Kosoulas, A. (Eds.). (2018). *Language teacher psychology*. Multilingual Matters.
- Mitsui, A., Donato, R., & Tucker, G. R. (2009). Introducing writing activities into the Japanese-as-a-foreign-language (JFL) classroom: A teacher-researcher collaboration. *NECTFL Review*, 64, 5–17.
- Pessoa, S., Hendry, H., Donato, R., Tucker, G. R., & Lee, H. (2007). Content-based instruction in the foreign language classroom: A discourse perspective. *Foreign Language Annals*, 40(1), 102–121.
- Shiveley, J., & Misco, T. (2015). Long-term impacts of short-term study abroad: Teacher perceptions of preservice study abroad experience. *Frontiers: The Interdisciplinary Journal of Study Abroad*, XXVI, 107–120.
- Tucker, G. R. (1974). *Innovative approaches to second-language teaching*. TESOL.
- Tucker, G. R. (1983). Developing a research agenda for second language educators. In J. E. Alatis, H. H. Stern, & P. Strevens (Eds.), *Georgetown University round table on languages and linguistics 1983* (pp. 386–394). Georgetown University Press.
- Tucker, G. R. (2000). Precision, elegance and simplicity: Perspectives on TESOL and art. *Journal of the Imagination in Language Learning*, 5, 24–26.
- Tucker, G. R. (Ed.). (1993). *Policy and practice in the education of culturally and linguistically diverse students: Views from language educators*. TESOL.
- Tucker, G. R., & Crandall, J. (1989). The integration of language and content instruction for language minority and language majority students. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1989* (pp. 39–50). Georgetown University Press.

- Tucker, G. R., & Donato, R. (1995). Developing a second language research component within a teacher education program. In J. E. Alatis F(Ed.), *Georgetown University round table on languages and linguistics 1995* (pp. 453–470). Georgetown University.
- Tucker, G. R., & Lambert, W. E. (1970). The effect on foreign language teachers of leadership training in a foreign setting. *Foreign Language Annals*, 4(1), 68–83.
- Walters, J. M., Garii, B., & Walters, T. (2009). Learning globally, teaching locally: Incorporating international exchange and intercultural learning into pre-service teacher training. *Intercultural Education*, 20(sup1), S151–S158.

Crossing the Disciplines: State of TESOL Teacher Education Programs in US Universities



Megumi Hamada and Ryan T. Miller 

Abstract The establishment of TESOL as a professional field in the 1960s was led by the Center for Applied Linguistics (CAL), where Dick Tucker played a crucial role in shaping its missions and strategies and deconstructing diverse boundaries for understanding the profession and supporting teachers (Crandall & Tucker, 1990; Tucker, 1993). Due to its origin, TESOL programs are traditionally closely tied with the fields of applied linguistics and second language acquisition, with an emphasis on language learning and teaching. Over the past few decades as the number of English learners in U.S. schools and the need for qualified teachers have increased, many TESOL programs have begun offering training and preparation for elementary and secondary school teachers. The curricula of such licensure programs are accredited by the Council for the Accreditation of Educator Preparation (CAEP) and structured according to TESOL standards such as the Standards for Initial TESOL Pre-K–12 Teacher Preparation Programs (TESOL International Association (TESOL) Standards for initial TESOL pre-K-12 teacher preparation programs. Alexandria, VA: Author, 2019). Although the standards integrate applied linguistics/SLA into the field of education, some licensure programs, compared with others, tend to have more emphasis on education and instruction than on learning and acquisition. In this chapter, drawing upon our own experience and professional engagement in TESOL, we evaluate the distinct focuses of TESOL programs in U.S. universities, in particular whether they are more applied linguistics/SLA-oriented or education-oriented, to provide insights into varied practices in TESOL teacher preparation and development in the U.S. In this evaluation, we consider factors such as home department (where the program is housed) and its faculty expertise, academic level (undergraduate vs. graduate), and licensure vs. non-licensure offerings. Based on the evaluation, we offer recommendations for twenty-first century TESOL programs that cross the disciplinary boundary between applied

M. Hamada (✉) · R. T. Miller
Department of English, Ball State University, Muncie, IN, USA
e-mail: mhamada@bsu.edu; rmill129@kent.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_6

linguistics/SLA and education by incorporating research and practices from both fields.

Keywords English learners · Teacher education · Knowledge base · Applied linguistics · TESOL

1 Introduction

The number of culturally and linguistically diverse students in U.S. schools has been steadily increasing for many years. Between the 2009–2010 and 2014–2015 school years, the percentage of English learners (ELs) increased in over half of states, with increases of more than 40% in five states (US Department of Education, 2018). Over the past 30 years, the number of ELs has increased from 3.6 million (representing 6% of all K-12 students) to over five million (over 10% of K-12 students) (National Center for Education Statistics, 2021; Tucker, 1981).

This on-going increase in the prevalence of ELs has led to an acute need for teachers who are adequately and appropriately prepared to teach ELs. Universities have responded to this need by offering a variety of EL teacher preparation programs, such as add-on programs that lead to endorsement for teaching ELs or content-area teacher education curricula that are infused with coursework on teaching culturally and linguistically diverse students. However, even with these efforts, the majority of states in the U.S. have not been able to staff an adequate number of qualified EL teachers at schools (Office of English Language Acquisition, 2021) and there are concerns that content-area teachers may still be unprepared to effectively teach ELs (García et al, 2010; Villegas et al., 2018).

In the U.S., certifications or licenses for K-12 teachers are issued by individual states, according to the qualification requirements designated for specific grade levels or subject areas, including those for EL teachers. There is also a national-level accreditation instituted by the Council for the Accreditation of Educator Preparation (CAEP; formerly the National Council for the Accreditation of Teacher Education, NCATE) using standards developed by TESOL International Association. First published in 2001 and revised in 2010 and 2019, the TESOL/CAEP standards outline the content, pedagogical knowledge, and skills that are necessary for EL teachers to have, including knowledge of language structure, sociocultural context, methods of instruction and assessment, and professionalism and leadership (TESOL, 2019).

Institutionally, EL teacher education programs are housed in a variety of departments, including linguistics, language, or education departments. This reflects the various knowledge bases that language teaching and language teacher education draw on. Traditionally, applied linguistics formed the core of language teacher education, as the goal of instruction was primarily the learning of language (Crandall, 2000). However, the work of many EL teachers now includes not only teaching

language but making content-area information accessible to ELs, requiring knowledge and methods of teaching from outside of language teaching. As a result, theory and practice of education more generally (i.e., areas of education outside of language teaching) have exerted increasing influence in pre-service and in-service language teacher education (Crandall, 2000). In this way, language teacher education involves the crossing of disciplinary boundaries between two different knowledge bases: applied linguistics and education.

The goal of this chapter was to describe the current state of EL teacher education in the U.S. by examining the EL teacher education programs accredited by CAEP. In particular, we investigated the extent to which EL teacher education programs cross boundaries in terms of the knowledge bases that they draw on and how this boundary crossing is related to (1) the institutional home of EL teacher education programs and (2) programs' alignment with the TESOL/CAEP standards.

2 ELs in K-12 Schools in the US

Whether they are born in the U.S. or another country, ELs grow up in an environment where a language other than English is used as the primary language. Depending on the amount of exposure to English and their primary language, some ELs are able to use basic communicative English but may struggle with academic English (Cummins, 1979). Some ELs are fluent and literate in their primary language but may be new to English with very little knowledge. A few decades ago, the term *limited English proficient* (LEP) student was more commonly used to refer to an EL. However, this term (LEP) induced a negative connotation and did not acknowledge students' primary language ability as an asset. Although *EL* has a more neutral connotation than LEP, more recently, an alternative term, *emergent bilingual*, has been recommended, in order to acknowledge ELs' primary language background and their developing bilingualism (García, 2009). As expressed in the term, "bilingual," ELs are expected to maintain and develop language proficiency in both their primary language and English, rather than English only. In this chapter, we use the term *EL* to refer to any student who is enrolled in U.S. K-12 schools and is in the process of developing proficiency in English as an additional language (Wright, 2015).

Compulsory education in the U.S. starts at kindergarten, typically at age 5, and continues through the 12th grade. Per federal guidelines, schools identify ELs by conducting a home language survey and English proficiency assessment (Office of English Language Acquisition, 2016). Children who are identified as ELs receive additional educational support throughout the grade levels until they achieve adequate English proficiency. Because the administration of such education support varies across states and local educational agencies, the U.S. Department of Education Office for Civil Rights and the U.S. Department of Justice's Civil Rights Division published a "Dear Colleague" letter in 2015, pointing out legal mandates regarding EL education (Lhamon & Gupta, 2015). Under Title VI of the Civil Rights Act of

1964, state and local educational agencies are obligated to offer support to ensure that ELs can meaningfully participate in education programs and services. The Equal Educational Opportunities Act of 1974 also requires public schools and state educational agencies to take actions to overcome language barriers that impede equal participation by ELs and their families in their educational programs.

3 EL Teacher Education

EL teachers play an integral role in providing effective EL education at schools. Currently, there is no nation-wide certification requirement for EL teachers, although there are federal grants available for state and local educational agencies to offer EL-related professional development programs for teachers. Nevertheless, many states require EL teachers to be officially certified to teach at schools. According to the Education Commission of the States (2020a), 26 states require EL teachers to hold a certification or endorsement in ESL or bilingual education as a state statute or regulation. For instance, in the state of California, the Commission on Teacher Credentialing issues authorizations for teachers who provide specially designed content instruction delivered in English, content instruction delivered primarily in a primary (home) language, or instruction for English language development (Cal. Educ. Code § 44253.2 et.seq). Furthermore, California requires any content-area teacher who has one or more EL in their classroom to have a certificate or authorization (Cal. Educ. Code § 44253.7). In contrast, for example, in the state of Indiana, an EL certification requirement is not specified in statute or regulation, yet individual schools may require such a certification. The state assures that there are pre-service and in-service training programs for persons serving non-English dominant students as educational personnel (Ind. Code Ann. § 20–30–9–6).

EL teacher education has its roots in language teacher education, particularly the education of teachers of English as a second or foreign language (ESL/EFL), where the primary goal of instruction has typically been language learning. This focus on language is reflected in the teaching methods that are typically covered in undergraduate and graduate ESL/EFL methods courses (see, e.g., Larsen-Freeman & Anderson, 2011). For example, the grammar-translation method focuses on grammar learning through translation of literature in a second language, the audiolingual method uses repetition and memorization of dialogues as a tool for learning language patterns, and communicative language teaching focuses on speech acts and learning language for everyday communication (Larsen-Freeman & Anderson, 2011). These and other prominent ESL/EFL teaching methods draw largely on concepts and theories from applied linguistics and second language acquisition, which for many years formed the core of ESL/EFL teacher education (Crandall, 2000).

However, EL teachers in the U.S. K-12 context have unique responsibilities that differ from traditional ESL/EFL teachers who are primarily language teachers. EL teachers in U.S. schools need to be able to support students' knowledge and skill development in both language (linguistic and sociopragmatic) and content areas

(academic subjects) (Fradd & Lee, 1998). This content-based approach is grounded in sociocultural theory (Vygotsky, 1978), which argues that language, cognition, and learning are interconnected. Children learn by exercising cognitive thinking processes, which involves language. At schools, ELs are expected to not only learn the English language, but to use English to develop new knowledge in academic subjects. In other words, for ELs, language is both content (as linguistic knowledge and skills) and a tool for learning academic subjects (Lucas, 2010).

4 EL Programs and Integration of Language and Content

Responding to the need for both language and content learning, EL instruction in U.S. schools adopts content-based language teaching approaches. Content-based language teaching can be defined as “an integrated approach to language instruction drawing topics, texts, and tasks from content or subject matter classes, but focusing on the cognitive, academic language skills required to participate effectively in content instruction” (Crandall & Tucker, 1990, p. 83). Content-based and language-based approaches are two ends of the continuum regarding the role of language in instruction (Met, 1999). Content-based approaches integrate language skills into the teaching of content knowledge and have as their goal the teaching of both content and language, whereas language-based approaches focus on teaching language skills with some aid from context and have as their primary goal the teaching of language. Thus, an EL teacher’s role is different from that of traditional ESL/EFL teachers who are primarily language teachers.

The integration of language and content learning is not a new concept, of course. The need for such integration was one of the main findings of early bilingual education research, such as the 12-year longitudinal study of French immersion conducted by Lambert and Tucker (1972). This and subsequent research has found that integrating language instruction with content instruction (rather than teaching language and content separately) can facilitate second language acquisition while simultaneously resulting in content learning that is similar to that of students taught in their native language (Tucker & Crandall, 1989).

EL teachers’ roles vary widely depending on the needs of schools and communities. In some models of instruction, the EL teacher has primary responsibility for instruction while in others the content-area teacher has primary responsibility or the two share responsibility (Crandall & Tucker, 1989). Table 1 summarizes EL program models that are commonly offered in U.S. schools, based on Lindahl and Baecher (2019). The programs include submersion, ESL pull-out, co-teaching (or ESL push-in), sheltered English or structured immersion, transitional or maintenance bilingual, and dual immersion bilingual programs. These programs differ in terms of setting and instructional language used; however, all involve some degree of integration of content and language.

Table 1 EL program models

Program	Characteristics
Submersion	ELs are educated in the content-area classroom with their English-speaking peers. Instruction is in English.
ESL pull-out	ELs are educated in a small group or one-on-one setting, separated from the content-area classroom and their English-speaking peers. Instruction is in English.
Co-teaching (or ESL push-in)	ELs are educated in the content-area classroom. An EL teacher co-teaches with the content-area teacher. Instruction is in English.
Newcomer	ELs are educated in a school or program with all other ELs who have been in the US less than 2 years. All subjects are taught with sheltered content techniques. Instruction is in English
Sheltered English or structured immersion	ELs are educated in the content-area classroom with their English-speaking peers. The teacher uses techniques for “sheltering” ELs that specifically foster language development. Instruction is in English.
Transitional or maintenance bilingual program	ELs of one primary language group in the elementary grades (K-3) are educated apart from their English-speaking peers or L2 learners from other backgrounds. Instruction is in English and their primary language.
Dual immersion bilingual programs	ELs are educated in a bilingual environment with their English-speaking peers, usually grades K-6. Instruction is in English and their primary language.

4.1 *EL Teacher Knowledge Bases*

In order to provide EL teachers with the knowledge and skills necessary to integrate language and content, EL teacher education must draw on diverse knowledge bases (Fradd & Lee, 1998; Mullock, 2006). For example, Day (1993) describes second language teaching (in general, not just EL teaching) as drawing on both domain-specific knowledge from TESOL and applied linguistics as well as general pedagogic knowledge from education more broadly. Domain-specific knowledge includes content knowledge (knowledge of the English language), pedagogical content knowledge (the specific ways of teaching second or foreign languages), and support knowledge (knowledge from disciplines that support language teaching such as psycholinguistics, sociolinguistics, and second language acquisition), while general pedagogic knowledge includes knowledge of general teaching strategies, beliefs, and practices, regardless of the specific subject being taught (such as classroom management strategies, motivational strategies, and pedagogical decision making strategies).

Existing EL teacher education frameworks also encompass multiple knowledge bases as core components. For instance, Lucas and her colleagues (Lucas & Villegas, 2010, 2013; Lucas et al., 2008) propose the following components in their EL teacher education framework: (1) the orientation of linguistically responsive teachers and (2) pedagogical knowledge and skills of linguistically responsive teachers. The first category addresses sociolinguistic consciousness and advocacy for ELs and linguistic diversity. The second category addresses application of second language learning theories to instructional strategies.

Coady et al. (2015) suggest that EL teacher education should be built on three dimensions: teachers' background and experience, teachers' knowledge about ELs, and teachers' knowledge of teaching and learning for ELs. The third dimension, which is about knowledge related to instruction, is further categorized into three components: (1) linguistics, (2) culture and SLA, and (3) instructional practices. The first component includes teachers' knowledge of the structure of languages. The second component includes understanding of the role of culture in ELs' learning of English language and academic content. The third component refers to the ability to provide differentiated instruction to meet ELs' various language learning needs. The Coady et al. framework describes more explicitly the connection between English language learning and content learning.

The framework by TESOL International Association (2019), which is used by CAEP for EL teacher education program accreditation, describes the standards for Pre-K-12 grade teacher preparation programs in the following five domains: (1) knowledge about language, (2) ELs in the sociocultural context, (3) planning and implementing instruction, (4) assessment and evaluation, and (5) professionalism and leadership. This framework is designed to prepare teacher candidates to effectively serve linguistically and culturally diverse students at U.S. K-12 schools. The first domain addresses the knowledge of linguistic systems and second language acquisition processes. The second domain addresses the role of identity and socio-cultural context in supporting ELs and their families. The third domain addresses the knowledge of culturally and linguistically supportive lessons to support the learning of language and content. The fourth domain addresses knowledge of various tools to assess language development. The fifth domain addresses the knowledge of effective collaboration with other educators and personal growth as reflective teachers. Similar to the framework by Coady et al. (2015), the TESOL standards aim to prepare EL teachers to be competent in teaching language and content for ELs with various proficiency levels.

4.2 Challenges in Teaching both Language and Content

Although the goal of EL teacher education programs is to prepare teacher candidates to become able to teach language and content, researchers have warned that teachers may face difficulty identifying language features specific to academic content (e.g., Schleppegrell, 2001, 2007; Turkan et al., 2014). For instance, in teaching the Pythagorean Theorem (e.g., $c^2 = 25$, $c = 5$), teachers would first need to notice that "taking the square root of 25" and "squaring 5" are the key linguistic features and explain to ELs that those two expressions refer to inverse operations. Then, the teachers need to model for the ELs how to use the expressions orally and in writing in the context of mathematics. Each academic discipline has its own unique context in which a linguistic register operates. Unless EL teachers are knowledgeable about such discipline-specific linguistic demands, including the lexicon, morphosyntax, and pragmatics of the discipline, it is not possible to teach language and content effectively in an integrated manner.

Content-language integrated models, such as the Cognitive Academic Language Learning Approach (CALLA; Chamot & O'Malley, 1994) and the Sheltered Instruction Observation Protocol (SIOP; Echevarría, et al., 1999), have been developed in response to the instructional needs specific to academic subjects. Nevertheless, such models are primarily for the purpose of shaping instructional approaches, rather than specifying language features. Accordingly, EL teachers need to be able to identify specific linguistic features (e.g., vocabulary, grammar, etc.) that are necessary for teaching specific academic concepts, based on EL teachers' own experience and discretion. EL teachers who are primarily trained to be experts in English language, including linguistics and second language acquisition processes, are tasked with a responsibility that may be beyond their capacity. For instance, Kong (2009) found a stark difference between teachers primarily trained in content (science) instruction and teachers primarily trained as language teachers in eighth grade classrooms. The language-trained teachers were unable to provide in-depth content knowledge during the class, which resulted in a lesson in which the students discussed content they already knew using language they already knew.

Similarly, Tigert and Peercy (2017) found that language-trained teachers were not adequately prepared for teaching content at K-12 schools. In their qualitative study, four teacher candidates in graduate-level TESOL programs with K-12 ESL teaching credentials were tracked over one semester, during which they taught as interns in content classes at secondary schools. Because none of the teachers had an adequate background in the content areas, they struggled to understand the content when they planned lessons. The analysis of the lesson observations and interviews demonstrated that the teacher candidates were well trained to teach English, with their solid understanding of linguistics and second language acquisition, but they found it difficult to teach concepts in the content areas, due to a lack of content knowledge. These findings underscore the challenge of EL teacher education programs. EL teachers are expected to be experts in TESOL with a wealth of knowledge in language, diversity, and second language learning processes. However, when they start teaching, they are also expected to understand how that TESOL-specific knowledge base can be applied to the teaching of academic content, such as mathematics, social studies, and science, and assessing students' conceptual understandings. This clearly suggests that the traditional approach in TESOL of primarily training language teachers may not be sufficient for EL teachers without content-area knowledge or certification, and it underscores the necessity of crossing this boundary between language teaching and content instruction in EL teacher education.

5 EL Teacher Education for all Teachers

Although EL teachers may assist content teachers as resource staff or co-teach with the content teachers, it is not likely that EL teachers can always be present in every classroom where there is an EL. EL teachers may be able to teach ELs in pull-out

sessions, but it is usually not possible to meet with them every day. Because of the shortage of EL teachers, in reality, for most of the school day, ELs are placed in classrooms taught by content-area teachers who may or may not have any specialized training in TESOL. Studies suggest that content-area teachers' lack of TESOL training can lead to a range of misconceptions about EL teaching, such as having stereotypes and biases toward minority or diverse students (Kumar & Hamer, 2012) and mis-placing ELs in special education classrooms (Stein, 2011). The misconception also affects content-area teachers' teaching approaches. Teachers tend to view EL-specific instruction as equivalent to that for any other diverse students (e.g., approaches in multicultural education) and consider good teaching of ELs as simply the same as good teaching for native-speaking students (Harper & de Jong, 2004; Pass & Mantero, 2009).

Responding to the growing need for all teachers to be equipped with TESOL-specific training, federal guidance encourages states to provide personnel to effectively facilitate EL programs, including content-area teachers who have received training to support ELs in their classroom (Office of English Language Acquisition, 2016). According to data from the Education Commission of States (2020b), 27 states require or provide TESOL training for all teachers, although the extent of the training varies greatly. Only a few states require an EL certification or endorsement (e.g., California), and the majority require TESOL training by means of the inclusion of TESOL in content-area teacher education programs, the inclusion of TESOL in the states' teacher qualification standards, or TESOL-specific professional development. Responding to the need, many teacher education programs have taken action to include TESOL-specific training in the existing curriculum. There are four commonly used approaches to implement such changes: (1) add a course, (2) modify existing courses and fieldwork to infuse attention to teaching ELs, (3) modify prerequisites, and (4) add a minor or additional certification (Lucas & Villegas, 2010).

For instance, de Jong and Naranjo (2019) report the efforts and struggles in teacher education programs in Florida, where elementary preservice teachers are required to have an EL endorsement. This requirement can be met by completing five courses in TESOL (through an endorsement program) or through an infused model with a minimum of two TESOL courses with additional general education courses that include EL-specific knowledge (FDOE, 2001). The TESOL courses are taught by faculty with expertise in TESOL, applied linguistics, bilingual education, or related fields, while the infused courses are taught by general education faculty who have completed 45 hours of professional development in TESOL. Based on an analysis of the infused courses, de Jong and Naranjo concluded that the general education faculty would need more professional development to effectively infuse their courses. Likewise, the data from de Jong et al. (2018) demonstrated that 74% of TESOL faculty considered the general education faculty at their institutions to be either not prepared or not well prepared to infuse EL knowledge and skills into their courses even after the state-required professional development.

Although there are findings that recognize the benefit of the infusion approach on teacher candidates' perception and instruction (e.g., Coady et al., 2011; Hutchinson, 2012; Lavery et al., 2019), teacher education faculty need to more strongly commit

to making changes to their existing courses and pedagogy in order to make infused courses more meaningful to teacher candidates (e.g., Costa et al., 2005). In addition, Baecher and Jewkes (2014) argue that collaboration between general education and TESOL faculty is crucial. The researchers implemented a semester-long collaboration between an early childhood education (ECE) class and a TESOL practicum class in which the ECE and TESOL faculty and their student-teachers collaborated in joint class sessions and discussed EL-specific strategies using sample lesson videos. The student-teachers from both classes expressed the benefit of such collaboration. In particular, the collaboration had a clear impact on the ECE student-teachers' perception about ELs and EL pedagogy.

6 Boundary Crossing in EL Teacher Education and the Present Study

From the research described earlier, it is clear that effective EL teachers necessarily cross boundaries between disciplinary knowledge bases, particularly applied linguistics and language teaching on one hand, and general and content-area education on the other. We see that teachers whose preparation focused on one without the other may be underprepared to integrate content and language for teaching ELs in U.S. schools today (Coady et al., 2015; de Jong & Naranjo, 2019; de Jong et al., 2018; Kong, 2009; Tigert & Peercy, 2017). It is, thus, important to investigate the extent to which EL teacher education programs facilitate such boundary crossing.

In order to examine boundary crossing in EL teacher education programs, we analyzed curricula/coursework of EL teacher education programs. To our knowledge, there have not been any studies that evaluated EL teacher education programs specifically, although several studies have evaluated curricula in MA TESOL programs. For instance, Ramanathan et al. (2001) investigated the cultures of two MA TESOL programs in universities in different parts of the U.S. and found that each program's identity and coursework was influenced by the culture and priorities of its home department. More recently, Stapleton and Shao (2018) conducted a curriculum survey of MA TESOL programs in 16 countries, including the U.S. They categorized the courses offered in the programs according to 15 knowledge bases and found that courses in three knowledge bases, teaching methods, linguistics, and SLA theories, were the most frequently covered in the programs. However, they found that, overall, the programs varied widely in their coursework and knowledge bases. In particular, the practicum/internship knowledge base appeared to be more popular among MA TESOL programs in the U.S. that also offered a K-12 EL license.

The national data indicate that ELs often underperform in academic achievement and are more likely to drop out of school than their native English-speaking peers (Sheng et al., 2011). To ensure all ELs succeed at school, EL teachers need to be able to provide effective instruction of both language and content. Based on the literature reviewed earlier, this study focuses on two knowledge bases that are integral in EL teacher education programs: applied linguistics and education. The applied linguistics

knowledge base refers to a repertoire of knowledge typically included in a TESOL or second/foreign language teaching program, such as linguistics, sociolinguistics, second language acquisition, and language teaching methodology. The education knowledge base refers to a repertoire of knowledge necessary to teach academic subjects in content areas. The question that motivated this study regards the extent of coverage and integration of the two knowledge bases in EL teacher education programs.

7 Method

7.1 Data Collection

We modeled our data collection and coding procedures after Stapleton and Shao (2018). The basic procedures were to collect curriculum information for EL teacher education programs and code the courses according to the knowledge base of each course. As a way to sample EL teacher education programs for this study, we collected information about the programs accredited by the Council for the Accreditation of Educator Preparation (CAEP) based on the standards developed by TESOL International Association. At the time of data collection in April of 2021, there were 103 programs listed on the CAEP website (<http://caepnet.org/provider-search>). For each program, two research assistants visited the university website and recorded the home department where the program was housed and the curriculum (total required credit hours and course titles and their credit hours). Unfortunately, ten of the programs did not have information available on their university website. Accordingly, the total number of programs submitted for coding and analysis was 93 (18 endorsement, 17 post-baccalaureate, 3 post-masters, 12 baccalaureate, and 43 masters programs).

A few programs listed different track options (e.g., an elementary school track vs. a secondary school track). For those programs, we chose only the first track shown on the program website to avoid duplication. The average credit hours for the program types were 17.50 for endorsement, 18.12 for post-baccalaureate, 20.00 for post-masters, 74.67 for baccalaureate, and 35.09 for masters. Because the endorsement, post-baccalaureate, and post-masters programs were all non-degree programs and had similar credit hour requirements, we decided to group them together in the subsequent coding and analysis. Table 2 summarizes the number of programs coded

Table 2 Number of programs by home department

	Masters	Endorsement & Post-Bacc/Masters	Baccalaureate	Total
Education	31	25	6	62
Language	6	11	6	23
Other	6	2	0	8
Total	43	38	12	93

and analyzed in this study, categorized by home department. The programs housed in “education” departments included departments, colleges, and schools of education and related disciplines (e.g., teaching and learning, curriculum and instruction). The education departments were by far the largest group with a total of 62 programs. The programs housed in “language” departments included departments and programs in linguistics, world languages, language education, English, or TESOL, comprising a total of 23 programs. The programs housed in “other” departments included those that did not belong in either the education or language category (e.g., literacy, graduate school), comprising only 8 programs.

7.2 Data Coding

All of the credit-bearing courses were coded for each program, totaling 1159 courses. The mean number of courses were 13.65 ($SD = 4.35$) for the Masters programs; 6.63 ($SD = 1.85$) for the endorsement, post-baccalaureate, and post-masters programs; and 26.67 ($SD = 5.63$) for the baccalaureate programs. Some programs included electives in the curriculum, in which students needed to complete a required number of courses from a set of course options. Because course choice could affect results, instead of making arbitrary choices, we decided to code all of the elective options in order to capture the range of knowledge covered in the curriculum.

Each course was coded in two categories, knowledge base (applied linguistics, education, both, or other) and the domains in the TESOL/CAEP standards. As for knowledge base, the courses coded as “applied linguistics” (AL) were the courses in applied linguistics or TESOL without specific reference to K-12 education or content-area instruction. Some example courses were SLA and Teaching, SLA and Assessment, Introduction to Linguistics, Bilingualism, Analysis and Structure of English, and Methodology of TESOL. The courses coded as “education” (ED) were courses in education without specific reference to TESOL, ELs, or bilinguals. Some example courses were Curriculum Theory and Instruction, Cultural Diversity and Education, Assessment of Learning, and Ethical and Moral Foundations of Educational Leadership. The courses coded as “both applied linguistics and education” (Both) were courses that integrated the applied linguistics/TESOL-specific knowledge into K-12 education or content-area instruction. It is these courses (those coded as Both) that we see as demonstrating the kind of boundary crossing that is most necessary in EL teacher education in the U.S. Some example courses were Language Arts and ESL Instruction, K-12 ESL Curriculum and Materials across the Content Areas, Assessment of Bilingual Students, Linguistics for Language Teachers, and Supervised Student Teaching in TESOL Grades K-6. The courses coded as “other” were courses that were outside of the TESOL or education disciplines (e.g., Political and Cultural Geography, Introduction to Psychology).

The courses coded as either AL, ED, or Both were also coded into the five domains of the Standards for Pre-K-12 Teacher Preparation Programs developed by

TESOL International Association (2019). Coding was based on the core domain sought in each standard as follows: (S1) linguistic systems and SLA theories/processes, (S2) diversity and sociocultural roles, (S3) teaching methods, curriculum, and materials, (S4) assessment and evaluation, (S5) supervised teaching, EL education policies, and leadership. The courses that did not belong to any of the standards (e.g., Introduction to Special Education, Philosophy of Education) were coded as “not applicable (n/a).”

For the course titles that were unclear (e.g., Culture), we looked up the course description on the program website to determine the coding. After initial discussion of the coding criteria, both authors coded 21 programs (3 endorsement, 3 post-baccalaureate, 1 post-masters, 4 baccalaureate, and 10 masters programs) independently and compared the coding results. At this initial stage, we were able to obtain 75% agreement. We discussed and resolved each of the courses where there were coding discrepancies and established more detailed criteria. The remaining programs were coded by the first author using the updated version of the criteria.

8 Results and Discussion

8.1 Knowledge Base

Table 3 summarizes the mean percentages of the courses within each knowledge base for each program type. The knowledge bases for the masters programs and for the endorsement and post-baccalaureate/masters seemed to follow a similar pattern, with AL and Both covering a large proportion of the courses (together accounting for 73.34% in the master’s programs and 89.24% in the endorsement and post-bacc/master’s programs). The small proportion of ED courses may reflect the nature of add-on programs, which are often designed for those who already have K-12 licensure in another subject and need more EL-specific coursework. It is also worth noting that the Both category represented the largest portion of courses in both the masters programs and add-on programs. This may reflect efforts to integrate EL-specific content into education courses. However, we see that the knowledge bases for the baccalaureate programs were more evenly distributed between AL and ED (37.83% and 31.89%, respectively), with a smaller percentage of courses including both knowledge bases (approximately 20%) and other (approximately 10%).

Table 3 Mean percentages of courses by knowledge base and program type

	AL Knowledge	ED Knowledge	Both Knowledge	Other
Master’s	33.35 (20.41)	26.66 (17.09)	39.99 (17.88)	0
Endorsement & Post-B/M	43.58 (22.24)	10.85 (18.67)	45.66 (20.91)	0.38 (2.32)
Baccalaureate	37.83 (14.06)	31.89 (15.06)	19.12 (12.64)	11.16 (7.26)

Note. Numbers in parentheses are standard deviations

The distribution of the AL and ED knowledge bases may be due to the influence of home department, which we examine in the next section.

8.2 Home Department and Knowledge Base

Table 4 provides a breakdown of the mean percentages of the curriculums that draw on each knowledge base according to the program type and home department. For the masters programs, the course coverage in programs housed in education departments was almost identical between the AL and ED knowledge bases (approximately 30% each) with about 40% coverage in the Both knowledge base. However, the course coverage of programs housed in language departments leaned more towards the AL knowledge base (52.21%), followed by 30% for the Both knowledge base and only 15% for the ED knowledge base. Similarly, in the endorsement and post-baccalaureate/masters programs, language departments had more than 60% of the courses in the AL knowledge base, with only 10% in ED and 25% in Both, while education departments had approximately 40% in AL and 50% in Both knowledge bases. Collectively, these data seem to suggest that among EL teacher education programs offered by language departments, more than half of the curriculum draws on the traditional applied linguistics and TESOL knowledge base with less exposure to general education knowledge or integration of K-12 EL teaching contexts. On the other hand, the fact that programs housed in education departments had a higher proportion of courses that incorporate both AL and ED knowledge bases may suggest that more courses in these programs are purposely designed to cross the boundary between these two disciplines by integrating applied linguistics and TESOL-specific knowledge into K-12 EL contexts in order to prepare EL teachers to integrate language and content instruction. Examples of such course titles include K-12 Bilingual and TESOL Teaching Practices and Assessment in the Content Areas, Learning Content Through Language in Multilingual Classrooms, and Infusing Content Language Instruction into TESOL/Bilingual Programs.

For the baccalaureate programs, the percentages of the combined (Both) knowledge base are only about 10% and 25% in education departments and language departments, respectively. The percentage distribution between the ED and AL

Table 4 Mean percentages of courses by knowledge base and home department

	Department	AL Knowledge	ED Knowledge	Both Knowledge
Masters	Education	29.74 (19.36)	29.57 (17.92)	40.68 (17.58)
	Language	52.21 (21.82)	14.71 (7.35)	33.08 (22.70)
Endorsement &	Education	36.19 (16.81)	11.80 (15.69)	52.49 (16.84)
Post-Bacc/masters	Language	62.86 (23.45)	10.75 (25.75)	25.73 (15.64)
Baccalaureate	Education	30.03 (11.42)	42.27 (13.96)	13.32 (9.65)
	Language	45.63 (12.59)	21.50 (6.71)	24.92 (13.34)

Note. Numbers in parentheses are standard deviations

knowledge bases seemed to follow the home department, with programs in education departments having a higher percentage in the ED knowledge base (42.27%) than the AL knowledge base, whereas those in language departments had a higher percentage in the AL knowledge base (45.63%) than the ED knowledge base. This may suggest that baccalaureate programs tend to focus more on developing a foundational knowledge base in the department's field, with less emphasis on integrating the knowledge bases.

Overall, these findings suggest that home department has a substantial influence on the distribution of courses in a program. This echoes the findings of Ramanathan et al. (2001), who compared two MA TESOL programs in universities in different parts of the U.S. and found that each program's identity and coursework were influenced by its home department. They found that each of the programs conformed to the ideologies prevalent in the department. More specifically, in that study, a program that was housed in a linguistics department was oriented toward linguistic structure in order to fit into the department, and a program housed in an English department focused less on pedagogy in order to fit into its home department. In the present study, we see a similar pattern with language departments and education departments.

8.3 Knowledge Base and TESOL Standards

The mean percentages of courses by knowledge base and TESOL standard domain are shown in Table 5 for programs in education departments and in Table 6 for those in language departments. Note that the percentages in the tables include only the programs that had courses coded for either AL, ED, or Both knowledge bases. Some programs included no courses in one knowledge base, and, consequently, the total percentages are lower than 100% in some of the knowledge bases.

Standard 1 addresses teachers' knowledge about language and language acquisition, including knowledge of English language structures, English language use, second language acquisition and development, and language processes that help ELs acquire language specific to various content areas (TESOL, 2019). In general, programs drew heavily on the applied linguistics knowledge base to meet Standard

Table 5 Mean percentages of knowledge base in TESOL standards for the programs housed in education departments

	S1	S2	S3	S4	S 5	n/a
AL knowledge	38.48 (24.82)	6.56 (12.78)	33.55 (27.13)	10.45 (20.40)	1.77 (7.51)	5.97 (13.17)
ED knowledge	0.49 (2.42)	18.57 (32.17)	16.31 (21.68)	4.32 (10.52)	4.53 (14.42)	31.59 (35.44)
Both knowledge	17.55 (18.95)	6.30 (12.76)	33.63 (25.23)	9.35 (12.44)	30.71 (21.26)	0.85 (3.43)

Note. Numbers in parentheses are standard deviations

Table 6 Mean percentages of knowledge base in TESOL standards for the programs housed in language departments

	S1	S2	S3	S4	S5	n/a
AL knowledge	54.28 (20.80)	9.77 (11.24)	24.04 (16.00)	6.82 (7.83)	0	5.08 (9.72)
ED knowledge	0.87 (4.17)	12.80 (24.97)	12.01 (22.72)	6.24 (21.04)	2.17 (10.43)	22.43 (38.12)
Both knowledge	9.76 (17.75)	3.10 (9.59)	34.15 (32.61)	6.34 (21.68)	46.66 (35.58)	0

Note. Numbers in parentheses are standard deviations

1, regardless of the department where the program was housed (38.48% of the AL knowledge base in programs housed in education departments and 54.28% in programs in language departments). Nevertheless, the proportion of courses that met Standard 1 was noticeably higher in programs in language departments than those in education departments. This may suggest that the curricula offered by language departments tend to be more focused on linguistic systems and SLA theory, with less integration of applied linguistics/TESOL knowledge specifically in K-12 EL contexts. This again resonates with the findings of Ramanathan et al. (2001) regarding the influence of home department. Examples of AL knowledge-base courses that address Standard 1 are Structures of English, Second Language Acquisition, and Applied Linguistics. Programs housed in education departments tended to include more courses that integrated the AL and ED knowledge bases in order to address Standard 1; examples of these include Linguistics for PreK-12 ESOL Teachers, Applied Linguistics for Exceptional ELs/MLLs, and Language Structure and Analysis for ELL Teachers.

Standard 2 is related to knowledge of ELs' sociocultural context and includes knowledge of how personal, familial, cultural, social, and sociopolitical contexts affect ELs' learning (TESOL, 2019). In general, slightly more of the ED knowledge base was related to this standard than the AL knowledge base. Courses in the ED knowledge included those such as Development and Diversity; Race, Class, Gender, and Disability in American Education; and Culturally Sustaining Pedagogy. Courses drawing on the AL knowledge base to meet this standard were mainly courses on sociolinguistics, including titles such as Sociolinguistics and Mobility, Language in Society, and Socio-cultural Aspects of Language. Courses that combined the two knowledge bases to meet this standard included, for example, Cultural and Linguistic Diversity in Schools, Introduction to Culturally & Linguistically Diverse Learners, and Cultural Components of Bilingual and ESL Instruction.

Standard 3 has to do with planning and implementing instruction, including teachers' knowledge of teaching methods and "evidence-based, student-centered, developmentally appropriate interactive approaches" (TESOL, 2019, p. 9). A larger proportion of the AL knowledge base was dedicated to this standard than the ED knowledge base. AL knowledge base courses included, for example, Methods and Materials for Teaching ESL, Teaching Second Language Reading and Writing, and Methodology of TESOL. ED knowledge base courses included Multicultural

Education Methods and Materials, Technology and Teaching, and Digital Teaching and Learning in K-12 Schools. It is worth noting that a large proportion of the courses that integrated the AL and ED knowledge bases aligned with this standard, as seen in courses such as Methods and Materials of Teaching English as a Second Language through the Content Area Pre-K–12, Structured English Immersion and Sheltered English Content Instruction, and Infusing Content Language Instruction into TESOL/Bilingual Programs.

Standard 4 has to do with assessment and evaluation, including knowledge of classroom-based, standardized, and language proficiency assessments (TESOL, 2019) and represented a relatively small portion of each knowledge base. AL knowledge base courses that aligned with this standard were generally traditional second language assessment courses, such as Second Language Testing and Assessment, Assessment in TESOL, and Language Assessment. Courses in the ED knowledge base were more general educational measurement and assessment courses, such as Assessment of Learning or Classroom Assessment, or those that focused on other aspects of education than language, such as Assessment for Struggling Readers. Courses that combined the AL and ED knowledge bases included Assessment of Multilingual Learners, Testing and Evaluation of English Language Learners, and Testing, Assessment, and Evaluation in Bilingual and ESL Education. It is worth noting that this standard was addressed the least by programs.

Standard 5 addresses professionalism, leadership, and supervised teaching practice. This was also the standard where programs drew the most on the combined (Both) knowledge base (30.71% for programs housed in education departments and 46.66% for programs housed in language departments). The vast majority of programs addressed this standard through courses that involved supervised teaching practice, such as a teaching practicum or student teaching. However, there were some programs that also included courses that integrated the knowledge bases in addressing other aspects of this standard. For example, one program included a course on Teacher Leadership in TESOL and another included a course on ESL Leadership, Research, and Advocacy. However, such courses were very rare.

The courses in the ED knowledge base had higher percentages of not belonging to any of the TESOL standard domains (n/a): 31.59% for the education departments and 22.43% for the language departments. This may be because EL teacher education programs in education departments are often built on (and draw from) existing general education curricula. In such programs, courses related to multicultural education or diversity may be included in EL teacher education, even if they do not meet the TESOL standards. For example, some EL teacher education programs included special education courses or courses on generic multicultural education for minority students, such as students with Hispanic or African American backgrounds. Although such courses are related to diversity and multicultural education, they are not necessarily relevant for EL teachers and do not necessarily align with the standards.

As mentioned earlier, we also found that some programs included no courses in one knowledge base. Of the programs housed in education departments (Table 5), 3.23% included no courses in the AL knowledge base, 24.19% included no courses

in the ED knowledge base, and 1.61% included no courses in the combined Both knowledge base. Of the programs housed in language departments (Table 6), all included at least one course in the AL knowledge base and Both knowledge base, but 43.48% included no courses in the ED knowledge base. It is worth noting that a substantial portion of programs included no courses from the ED knowledge base, even programs that were housed in education departments.

9 Conclusions and Recommendations

In this study, we investigated the degree of boundary crossing between the fields of applied linguistics and general/content-area education in U.S. K-12 EL teacher education programs. To do this, we examined the degree to which programs of various types draw on the knowledge bases of applied linguistics and general/content-area education and, in particular, the degree to which programs include courses that integrate the two knowledge bases. Courses in the AL and ED knowledge bases are specific to their field, and teacher candidates whose training draws on the AL and ED knowledge bases separately would need to work on application on their own. For example, if the majority of their training draws on the AL knowledge base, teacher candidates would lack experience in applying this knowledge to K-12 EL contexts. On the other hand, if their training draws mainly on the ED knowledge base, teacher candidates would need to learn on their own about EL-specific aspects of instruction that are different from teaching native-speaking children or those introduced in courses on multicultural education more generally. Thus, to truly cross the boundaries between these disciplines, EL teacher education programs need courses that intentionally and purposefully integrate the two knowledge bases in order to provide an integrated approach.

One of the major findings was the influence of home department on the knowledge base that programs drew from, with programs housed in language departments drawing more heavily on the AL knowledge base and programs housed in education departments drawing more on the ED knowledge base. This finding is consistent with Ramanathan et al. (2001), who also found an influence of home department. Problems with such a situation have been pointed out by de Jong and Naranjo (2019) and de Jong et al. (2018), who found that some education faculty may receive training in TESOL/SLA as professional development but they are generally not well prepared to infuse their education courses with TESOL/EL-specific theories and practices. To address this issue, perhaps a more systematic change is needed in graduate programs for teacher educators. For example, including TESOL-specific courses and experience in doctoral programs in education would ensure that all teacher educators are fully prepared to infuse their content courses. Likewise, including K-12 EL-specific courses and experience in doctoral programs in applied linguistics and SLA would ensure TESOL faculty gain a better understanding of how TESOL-specific knowledge could be applied in K-12 contexts.

Another finding was that boundary-crossing courses (i.e., those that integrate both the AL and ED knowledge bases) were mainly limited to courses focused on planning instruction (Standard 3) and practical teaching experience (Standard 5). Boundary crossing was seldom the focus of courses that met other standards, such as courses on second language acquisition processes, sociocultural contexts, or, in particular, assessment (Standards 1, 2, and 4, respectively). We did find that a few programs did include such boundary-crossing courses, such as Linguistics for PreK-12 ESOL Teachers (Standard 1), Cultural Components of Bilingual and ESL Instruction (Standard 2), and Testing and Evaluation of English Language Learners (Standard 4). However, such integrated courses were very much in the minority and were largely confined to master's degree programs. EL teacher education programs need to be more proactive in developing courses that, by design, integrate the AL and ED knowledge bases.

K-12 schools in the U.S. have a pressing need for more qualified teachers to educate all ELs with academic excellence. In order to achieve this goal, teacher education programs need to be willing to cross boundaries between traditional disciplines. Crossing boundaries is likely to involve a transformation of teacher education programs with a curriculum that acknowledges EL-specific theories and practices, as recognized by Costa et al. (2005). Curricular changes may be carried out at the institutional level, but real program transformation seems to require changes at the individual level, including teacher education program faculty, administrators, and students.

References

- Baecher, L., & Jewkes, A. M. (2014). TESOL and early childhood collaborative inquiry: Joining forces and crossing boundaries. *Journal of Early Childhood Teacher Education, 35*, 39–53.
- Chamot, A. U., & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Addison Wesley Longman.
- Coady, M., Harper, C., & de Jong, E. (2011). From preservice to practice: Mainstream elementary teacher beliefs of preparation and efficacy with English language learners in the State of Florida. *Bilingual Research Journal, 34*, 223–239.
- Coady, M., Harper, C., & de Jong, E. (2015). Aiming for equity: Preparing mainstream teachers for inclusion or inclusive classrooms? *TESOL Quarterly, 50*(2), 340–368.
- Costa, J., McPhail, G., Smith, J., & Brisk, M. E. (2005). Faculty first: The challenge of infusing the teacher education curriculum with scholarship on English language learners. *Journal of Teacher Education, 56*, 104–118.
- Crandall, J. (2000). Language teacher education. *Annual Review of Applied Linguistics, 20*, 34–55.
- Crandall, J., & Tucker, G. R. (1990). Content-based language instruction in second and foreign languages. In S. Anivan (Ed.), *Language teaching methodology for the nineties* (pp. 83–96). SEAMO Regional Language Centre.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*(2), 222–251.
- Day, R. (1993). Models and the knowledge base of second language teacher education. *University of Hawai'i Working Papers in ESL, 11*, 1–13.

- de Jong, E., & Naranjo, C. (2019). General education teacher educators and English language learner teacher preparation: Infusion as curricular change. *The New Educator*, 15(4), 331–354.
- de Jong, E. J., Naranjo, C., Li, S., & Ouzia, A. (2018). Beyond compliance: ESL faculty's perspectives on preparing general education faculty for ESL infusion. *The Educational Forum*, 82, 174–190.
- Echevarria, J. J., Vogt, M. J., & Short, D. J. (1999). *Making content comprehensible for English language learners: The SIOP model* (1st ed.). Allyn and Bacon.
- Education Commission of States (ECS) (2020a). *50-State Comparison: English Learner Policies. Are EL-only instructors in K-12 required to hold a specialist certification or endorsement?* <https://reports.ecs.org/comparisons/50-state-comparison-english-learner-policies-11>
- Education Commission of States (ECS) (2020b). *50-State comparison: English Learner policies. What EL training or ongoing professional development is required of general classroom teachers?* <https://reports.ecs.org/comparisons/50-state-comparison-english-learner-policies-10>
- Florida Department of Education (FDOE). (2001). *Preparing Florida's teachers to work with limited English proficient students*. Bureau of Educator Recruitment and Professional Development.
- Fradd, S. H., & Lee, O. (1998). Development of a knowledge base for ESOL teacher education. *Teaching and Teacher Education*, 14, 761–773.
- García, O. (2009). Emergent bilinguals and TESOL: What's in a name? *TESOL Quarterly*, 43(2), 322–326.
- García, E., Arias, M. B., Harris Murri, N. J., & Serna, C. (2010). Developing responsive teachers: A challenge for a demographic reality. *Journal of Teacher Education*, 61(1–2), 132–142.
- Harper, C., & de Jong, E. (2004). Misconceptions about teaching English-language learners. *Journal of Adolescent and Adult Literacy*, 48, 152–162.
- Hutchinson, M. (2012). Bridging the gap: Preservice teachers and their knowledge of working with English language learners. *TESOL Journal*, 4(1), 25–54.
- Kong, S. (2009). Content-based instruction: What can we learn from content-trained teachers' and language-trained teachers' pedagogies? *The Canadian Modern Language Review*, 66(2), 233–267.
- Kumar, R., & Hamer, L. (2012). Preservice teachers' attitudes and beliefs toward student diversity and proposed instructional practices: A sequential design study. *Journal of Teacher Education*, 64(2), 162–177.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Newbury House.
- Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching* (3rd ed.). Oxford University Press.
- Lavery, M. R., Nutta, J., & Youngblood, A. (2019). Analyzing student learning gains to evaluate differentiated teacher preparation for fostering English learners' achievement in linguistically diverse classrooms. *Journal of Teacher Education*, 70(4), 372–387.
- Lhamon, C., & Gupta, V. (2015). *Dear colleague letter: English learner students and limited English proficient parents*. U.S. Department of Justice and U.S. Department of Education. Retrieved from <https://www2.ed.gov/about/offices/list/ocr/letters/colleague-el-201501.pdf>
- Lindahl, K., & Baecher, L. (2019). Preparing TESOL specialists for K-12 contexts. In L. C. de Oliveira (Ed.), *The handbook of TESOL in K-12* (pp. 357–369). Wiley Blackwell.
- Lucas, T. (2010). Language, schooling, and the preparation of teachers for linguistic diversity. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 3–17). Routledge.
- Lucas, T., & Villegas, A. M. (2010). The missing piece in teacher education: The preparation of linguistically responsive teachers. *National Society for the Study of Education*, 109(2), 297–318.
- Lucas, T., & Villegas, A. M. (2013). Preparing linguistically responsive teachers: Laying the foundation in preservice teacher education. *Theory into Practice*, 52, 98–103.

- Lucas, T., Villegas, A. M., & Freedson-Gonzalez, M. (2008). Linguistically responsive teacher education: Preparing classroom teachers to teach English language learners. *Journal of Teacher Education*, 59(4), 361–373.
- Met, M. (1999). *Content-based instruction: Defining terms, making decisions*. NFLC reports. The National Foreign Language Center.
- Mullock, B. (2006). The pedagogical knowledge base of four TESOL teachers. *The Modern Language Journal*, 90, 46–66.
- National Center for Education Statistics. (2021). *English language learners in public schools*. U.S. Department of Education. https://nces.ed.gov/programs/coe/pdf/2021/cgf_508c.pdf
- Office of English Language Acquisition. (2016). *English learner tool kit*. Rev. ed. U.S. Department of Education. <https://www2.ed.gov/about/offices/list/oela/english-learner-toolkit/index.html>
- Office of English Language Acquisition. (2021). *Teacher projections for English learners*. U.S. Department of Education. https://ncela.ed.gov/files/fast_facts/Del14-4ELTeacher%20Projections_6.11.2021_508.pdf
- Pass, C., & Mantero, M. (2009). (Un)covering the ideal: Investigating exemplary language arts teachers' beliefs and instruction of English language learners. *Critical Inquiry in Language Studies*, 6(4), 269–291.
- Ramanathan, V., Davies, C. E., & Schleppegrell, M. J. (2001). A naturalistic inquiry into the cultures of two divergent MA-TESOL programs: Implications for TESOL. *TESOL Quarterly*, 35(2), 297–305.
- Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431–459.
- Schleppegrell, M. J. (2007). The linguistic challenges of mathematics teaching and learning: A research review. *Reading and Writing Quarterly*, 23, 139–159.
- Sheng, Z., Sheng, Y., & Anderson, C. J. (2011). Dropping out of school among ELL students: Implications to schools and teacher education. *The Clearing House*, 84, 98–103.
- Stapleton, P., & Shao, Q. (2018). A worldwide survey of MATESOL programs in 2014: Patterns and perspectives. *Language Teaching Research*, 22(1), 10–28.
- Stein, J. C. (2011). The case for collaboration: Integrating information on English learners and special education in teacher preparation programs. *Multicultural Education*, 19(3), 35–40.
- TESOL International Association (TESOL). (2019). *Standards for initial TESOL pre-K-12 teacher preparation programs*. Author.
- Tigert, J. M., & Peercy, M. M. (2017). Preparing to teach both content and language: Four ESOL teacher candidates' experiences. *TESOL Journal*, 9(3), 542–556.
- Tucker, G. R. (1981). Language issues for the 1980s: Examples from the United States. *Canadian Modern Language Review*, 37, 441–446.
- Tucker, G. R. (Ed.) (1993). *Policy and practice in the education of Culturally and Linguistically Diverse students: View from language educators*. TESOL.
- Tucker, G. R., & Crandall, J. (1989). The integration of language and content instruction for language minority and language majority youngsters. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1989*. Georgetown University Press.
- Turkan, S., de Oliveira, L., Lee, O., & Phelps, G. (2014). Proposing a knowledge base for teaching academic content to English language learners: Disciplinary linguistic knowledge. *Teachers College Record*, 116(3), 1–30.
- U.S. Department of Education. (2018). *Our nation's English learners*. <https://www2.ed.gov/datas-tory/el-characteristics/index.html>
- Villegas, A. M., SaizdeLaMora, K., Martin, A. D., & Mills, T. (2018). Preparing future main-stream teachers to teach English language learners: A review of the empirical literature. *The Educational Forum*, 82, 138–155.
- Vygotsky, L. (1978). *Mind in society*. Cambridge University Press.
- Wright, W. E. (2015). *Foundations for teaching English language learners: Research, theory, policy, and practice* (second ed.). Caslon.

Computerized Mediation in the Instruction and Development of L2 Pragmatic Competence: A Dynamic Assessment Perspective



Tianyu Qin

Abstract Dynamic assessment (DA) breaks the traditional dichotomy between assessment and instruction by including mediation in assessment procedures ((Poehner *Dynamic assessment: A Vygotskian approach to understanding and promoting second language development*. Springer, Berlin, Germany, 2008)). The focus of DA is on how students or test-takers learn and change with mediation rather than what they can do alone. Traditionally, DA requires the presence of a person as the mediator, such as the teacher, to guide another individual's or student's learning. This human mediation is the focus of the majority of DA research. With the advancement of intelligent computing and its applications in language education, the boundaries between mediators and mediational tools can be crossed to facilitate language instruction and learning at any time and in any place. Against this backdrop, this study aims to expand DA research by applying pre-designed and standardized mediation via the computer, that is, computerized dynamic assessment (C-DA). It explores innovative and efficient mediational tools to facilitate second language (L2) development with a focus on L2 Chinese learners' pragmatic competence.

The study reported in this chapter centers on Chinese learners' ability to comprehend implicature (i.e., implied meaning) with computerized mediation through C-DA. The mediation is modeled based on pragmatics research, and it serves as a metacognitive tool. In particular, the mediation aims to help L2 Chinese learners 1) notice and pay attention to both literal and implied meanings, and 2) understand indirect speech acts in Chinese. Participants' test-taking processes were video recorded. Semi-structured interviews were conducted to let participants reflect upon their learning processes during C-DA. This chapter provides a case analysis of an intermediate-level Chinese learner's interactions with C-DA and how she used the computerized mediation to improve her performance during the test. The chapter ends with a discussion on the current situation of pragmatic instruction in language

T. Qin (✉)

Department of Modern & Classical Languages, University of North Georgia,
Dahlonega, GA, USA

e-mail: tianyu.qin@ung.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics
58, https://doi.org/10.1007/978-3-031-24078-2_7

149

classrooms, and talks about the potential of computerized mediation to facilitate the development of L2 pragmatic competence and to cross boundaries in learning environments and traditional teacher-student roles for ubiquitous learning of languages.

Keywords L2 Chinese · Pragmatics · Computerized mediation

1 Introduction

Through their *Educating for Global Awareness* project at Carnegie Mellon University in the 2000s, which investigated global literacy objectives in higher education, G. Richard Tucker (Dick) and his team identified major components of global literacy outcomes that are broadly conceived and disciplinarily meaningful. Among them, social/cultural competency is an essential aspect. Specifically, students with social/cultural competencies are able to “listen respectfully, recognizing differences in communication style and etiquette across cultures” (Nair et al., 2012, p.60). Along with the growing emphasis on intercultural competence in higher education, second language (L2) education has also shifted from emphasizing learners’ mastery of grammatical forms to increasing attention to their knowledge of socially appropriate language use, an essential aspect of Bachman and Palmer’s (1996) communicative competence model. Nowadays, the overarching goal of L2 education is to prepare students to become global professionals who are aware of cross-cultural differences and able to apply communicative strategies that “mediate across cultural and linguistic boundaries” (Taguchi & Roever, 2017, p.275).

To cross cultural and linguistic boundaries, L2 learners should be equipped with knowledge and skills in pragmatics. Interlanguage pragmatics is a field of second language acquisition (SLA) that centers on language use in social contexts as well as the development of L2 learners’ ability to comprehend and produce appropriate language in complex social interactions (LoCastro, 2011). Existing research supports the teachability of pragmatics (see Taguchi, 2015, for a review). In her review of research on instructed pragmatics, Taguchi (2011) pointed out that pragmatic competence was found to be enhanced through instruction under the noticing framework in SLA, explicit/implicit instruction, input processing, and skill acquisition and practice. With that being said, it is usually the researchers/teachers who provide instructional intervention to L2 learners, and most studies in this vein have been conducted in traditional experimental settings. To cross boundaries in traditional teacher-student roles for learning pragmatics as well as boundaries in learning environments, this chapter reports on a study that applied a computerized dynamic assessment (C-DA) instrument to play the role of “mediator” in order to assess and promote Chinese learners’ pragmatic competencies.

Drawing on Vygotskian sociocultural psychology, dynamic assessment (DA) integrates instructional intervention into an assessment program that aims to evaluate (1) learners’ solo performance and (2) what becomes possible with assistance (Poehner, 2008). Traditionally, DA has required the presence of a person as the mediator, such as a teacher, to guide another individual’s or student’s learning. This

human mediation is the focus of the majority of DA research. With the advancement of intelligent computing and its applications in language education, the boundaries between mediators or mediational tools can be crossed to facilitate language instruction and learning at any time and in any place. Against this backdrop, this study expanded DA research by applying pre-designed and standardized mediation via computer.

2 Background to the Study

2.1 Dynamic Assessment and Vygotsky's Zone of Proximal Development

DA has been applied in the fields of psychology and education to evaluate cognitive modifiability over the past decades (Haywood & Tzuriel, 2002). In other words, it is a form of assessment focused on the ability or readiness to learn rather than manifest performance. Although the concept of DA has been defined in various ways, all scholars perceive it as “an active intervention by examiners and assessment of the examinees’ response to intervention” (Haywood & Lidz, 2007, p.1). Unlike traditional assessment which is used to evaluate fully-developed human abilities, the DA approach has been implemented as a subset of interactive appraisal that centers on how learner abilities are formed through intervention (Vygotsky, 1978). The interactive negotiations between students and human mediators (tutors/teachers/assessors, etc.) during DA could provide insights into the obstacles to learner performance and provide instructional methods to overcome them through metacognitive mediation. It could also be used to assess the effects of the removal of these obstacles on subsequent performance (Haywood & Lidz, 2007).

DA is grounded in the Soviet psychologist L. S. Vygotsky’s writings on the Zone of Proximal Development (ZPD), which is defined as the distance between what an individual can achieve on his or her own, and what becomes possible with mediation (Vygotsky, 1978, 1986). Vygotsky claims that focusing on the evaluation of an individual’s current, already formed skills is past-oriented; in contrast, DA is development-oriented as it emphasizes a student’s potential skills that will emerge through mediation. Thus, the DA approach can also shed light on the student’s cognitive development potential (i.e., how far the individual can go with assistance).

2.2 Computerized Dynamic Assessment in L2 Research

The DA methodology has been extended to L2 settings in the past two decades to investigate language-related issues and to promote L2 development (Poehner, 2008). Most L2 DA research reported in the field has been conducted between an L2 learner and a human mediator in one-on-one settings, and the support has often been

offered flexibly and fine-tuned to better address learner needs. These studies investigated a wide range of L2 competencies, including speaking, writing, reading, listening, grammar, and pragmatics. Despite the fact that these small-scale qualitative DA studies have revealed predictive value for gauging learning potential, the evidence remains limited to specific contexts and small groups of participants (Qin & Zhang, 2018). To go beyond these highly contextualized settings, recent advances in L2 DA research have applied the computerized format of DA (C-DA), in which preprogrammed mediational prompts are integrated into the online testing procedures (Poehner & Lantolf, 2013; Poehner et al., 2015; Qin & van Compernelle, 2021).

Poehner and Lantolf (2013) and Poehner et al. (2015) developed reading and listening comprehension tests in L2 Chinese, French, and Russian. Their C-DA program was embedded with standardized, pre-scripted mediation during the assessment. The research team first worked with the targeted learners to identify potential difficulties on test items and subsequently used the information to design the standardized mediational prompts in their C-DA. The reading and listening C-DA applied a multiple-choice question (MCQ) format. Qin and van Compernelle (2021) extended this by using similar design principles to develop the first Chinese pragmatics listening C-DA test to measure L2 Chinese learners' pragmatic comprehension ability. A MCQ format was also applied. Following Aljaafreh and Lantolf (1994), mediation in both C-DA programs was designed to be graduated and contingent. Being graduated means the level of explicitness increased as needed, and being contingent means that mediation was available continuously throughout the test whenever students encounter difficulty. It is worth noting that both C-DA programs included transfer tasks, which consisted of items that are more challenging than regular items in the same test. The transfer tasks in DA aim to evaluate the extent to which learners can recontextualize their emerging abilities in various and more complex situations (Feuerstein et al., 1979; Poehner, 2007).

The two C-DA designs generated some common results with regards to learner development. The listening and reading C-DA projects reported that learners' performance on transfer tasks might not be predicted by their independent performance. Specifically, they found a group of students who improved significantly from unmediated to mediated performance, and this group did as well as the group of high achievers in the more challenging transfer tasks, indicating that students in general responded to the computerized mediation and new abilities emerged during C-DA. The pragmatics C-DA project reported participants' significant improvement across C-DA, suggesting that they were able to transfer the newly-formed pragmatic comprehension abilities to different contexts.

In sum, the two C-DA designs made initial attempts to apply the DA methodology at large scale. They both revealed how C-DA provided learning opportunities during the assessment and promoted learner development. However, one crucial issue is that learners' solo and mediated performance was not evaluated separately in existing C-DA work. That is because mediation was provided throughout the test, and unmediated and mediated performance was scored collaboratively (see Qin & Zhang, 2018 for a review). To make the two types of performance independent for the benefits of being easily comparable and interpretable, and, more importantly, to track learner change statistically during C-DA, this chapter reports on the design

and administration of a new C-DA instrument that crossed the existing boundaries by applying a “Pretest–Mediation–Near transfer–Far transfer” design.

2.3 Implicature Comprehension and L2 Instructed Pragmatics

Conversational implicature refers to the indirect, or implied meaning, derived from the context of an utterance and the shared knowledge of how conversations work (Grice, 1975). For instance, when being invited to an event that an individual has no intention to attend, they might come up with excuses as indirect refusals to turn the invitation down in order to avoid face-threatening acts. Existing L2 pragmatics research in implicature comprehension has focused almost exclusively on assessing whether or not L2 learners are able to comprehend implicatures. In addition, such research has mainly used nondynamic assessment tools; that is, only the manifest pragmatic competence is being examined. I argue that not being able to comprehend implied meanings on their own does not necessarily mean relevant pragmatic knowledge is not within learners’ L2 range of abilities. After all, implied language is commonly used in everyday conversations no matter what language is being used. It is predicted that new implicature comprehension abilities can emerge if L2 learners are given appropriate external support to make connections between implied meanings in L2 and their prior knowledge of how conversation works. This can be done efficiently through C-DA, but such research remains unexplored.

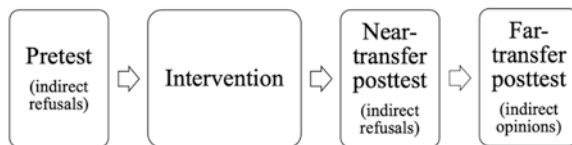
Theoretically speaking, the development of L2 pragmatics emphasizes the awareness or attention of learners (see Schmidt, 1993 for a review). Schmidt (1993) also claimed that it is necessary for language learners to pay attention to information such as “linguistic forms, functional meanings, and the relevant contextual features” (p. 35) to facilitate pragmatic learning. Accordingly, the C-DA design reported in this chapter includes instructional intervention that emphasizes literal utterance and contextual information in order to raise learner awareness of indirect meanings. The chapter reports on a case study that reveals how a focal participant’s pragmatic abilities emerged through C-DA. The information generated crosses boundaries in existing research on L2 pragmatic assessment and instruction and provides new insights into a combination of both fields that could potentially help language teachers teach pragmatics in an innovative and efficient way.

3 The Study

3.1 Context of the Research

The study reported in this chapter draws from data collected during a large research project on the design and administration of a C-DA instrument in assessing and promoting L2 Chinese learners’ implicature comprehension ability (Qin, 2018). Test items were originally designed by the researcher. Forty-five university learners

Fig. 1 C-DA procedures



of Chinese from two universities in the northeastern U.S. participated in the study in a computer lab. Participants were taking elementary or intermediate level Chinese classes at the time of data collection. Among them, 35 participated in a “test only” group in which learners were asked to complete the C-DA test and a background questionnaire (see [Appendix A](#)). Ten Chinese learners opted into a “test plus” group that completed the C-DA test, the background questionnaire, and an exit interview conducted by the researcher (for interview questions, see [Appendix B](#)). The exit interview aimed to probe learners’ test-taking processes as well as how their Chinese learning experience might have influenced their learning of implicatures. The “test plus” group’s interactions with C-DA and the exit interviews were recorded by a two-way video screen recording software called ScreenFlow that simultaneously captured the user screen and recorded streaming audio and video.

Following Zhang and van Compernelle’s (2016) three-stage “test-mediation-posttest” DA design, the current C-DA was designed as a four-part assessment: an independent pretest ($k = 12$), an adaptive intervention session with standardized prompts tailored to address the difficulties students encountered in the pretest, a near-transfer posttest ($k = 12$), and a far-transfer posttest ($k = 12$). Figure 1 displays the C-DA procedures. The pretest and near-transfer posttest evaluated L2 Chinese learners’ comprehension of the speech act of indirect refusals, and the far-transfer posttest evaluated their comprehension of a different speech act, indirect opinions, which are theoretically more challenging than indirect refusals (Taguchi, 2005; Taguchi et al., 2013). Test-takers completed the entire C-DA test in one sitting lasting approximately one hour.

Specifically, the indirect refusals involved excuses for refusing the other interlocutor’s requests, invitations, or suggestions. Example 1 is an indirect refusal of an invitation included in the test.

Example 1 Wenzhong invites his friend, Xiaomei, to go hiking during the weekend.

Xiaomei says,

“这个周末我要复习考试。”

“I will review for an exam during the weekend.”

The indirect opinions do not involve fixed discourse patterns and thus tend to be more idiosyncratic. Example 2 is an indirect positive opinion included in the test.

Example 2 Wenzhong asks his friend, Xiaomei, if she liked the movie she watched the

other day. Xiaomei says,

“那个电影你看了吗?我可以再跟你一起去看一次。”

“Did you watch that movie? I can go to watch it again with you.”

Comprehension questions in the pretest, near-transfer posttest, and far-transfer posttest applied a slide-ruler format. After listening to the audio file only once, a statement corresponding to the target indirect meaning was presented on the screen, and test-takers were asked to use a slide ruler to evaluate the extent to which they agree or disagree with the statement. The slide ruler was on a scale of -50 (strongly disagree) to $+50$ (strongly agree). The results generated from the slide ruler were converted to a scale of $0-100$ for the sake of data analysis. L2 Chinese learners' comprehension of individual items was calculated based on the corresponding threshold scores produced by a group of 20 native speakers of Chinese. For each item, 1 point was awarded if the learner's response was equal to or greater than the threshold score, while no points were awarded if the response was below the threshold score. For instance, if an item's threshold score is 75, learners whose responses were equal to or greater than 75 were awarded 1 point on that item, while learners whose responses were below 75 received 0 on that item.

In terms of the intervention session, it included standardized mediational prompts that were pre-designed in accordance with individual items. Each item was in a multiple-choice question (MCQ) format. Up to four prompts were available for each item. In line with the construct definition of implicature comprehension and DA methodology, the first prompt for each item simply allowed test-takers to try again, the second prompt mainly repeated the key content of the utterance (i.e., the locutionary act), the third prompt asked the intent of the utterance (i.e., illocutionary act), and last prompt provided the correct answer.

3.2 Focal Participant

The focal participant reported in this chapter, Christine, participated in the "test-plus" group. Christine was an 18-year-old freshman who was taking Intermediate Chinese I at her university at the time of data collection. She is an English-Gachi bilingual speaker. According to Christine, she had about 9 years of Chinese learning experience in total. Before university, she learned Chinese in elementary school (2–3 years), middle school (2–3 years), and high school (4 years consecutively). Her initial motivation to learn Chinese was that her elder brother started learning the language. She then became interested in Chinese language and culture and made Chinese-speaking friends in high school. Christine took a placement test at her university and was placed into the Intermediate Chinese I course. She did not report having visited a Chinese-speaking country before, but she was planning to participate in her university's summer study abroad program in Shanghai the following summer.

3.3 *The Genetic Method*

Following Lantolf and Poehner (2010), data in this chapter were analyzed with a genetic approach, underscoring how the C-DA instrument assessed and promoted the growth of Christine's emerging implicature comprehension ability. Genetic analysis is the overarching methodological approach in Vygotskian research (van Compernelle, 2014). It aims to trace how learners' intellectual functions begin and change over time regardless of the ontological and epistemological positions held by researchers (Wertsch, 1985). In Vygotsky's theoretical approach and method of experimentation, experimental results are both quantitative and qualitative in nature: "detailed descriptions, based on careful observation, will constitute an important part of experimental findings" (Vygotsky, 1978, p. 14). Therefore, in Vygotskian research, qualitative and quantitative methods, or a combination of both, are routinely used to trace the genesis of learner capacities. Methodologically, the genetic method crosses the traditional boundaries between SLA approaches (e.g., experimental instruments, qualitative coding), and provides an alternative for investigating the development of human mental behavior at a metatheoretical level (Lantolf & Thorne, 2006; van Compernelle, 2015).

4 **Emerging Implicature Comprehension Ability Mediated by C-DA**

This section presents a case analysis of Christine's performance during C-DA. Following the genetic method, the analysis investigated quantitative end products of the C-DA test (i.e., test scores) and qualitative discourse analysis on Christine's video-recorded test-taking processes and exit interview. Different sources of data were included to uncover Christine's development mediated by the computerized test.

4.1 *Quantitative Test Scores*

The quantitative scores Christine received during C-DA are summarized in Table 1. Her scores improved from the pretest to the near-transfer posttest, indicating that the intervention session supported her in comprehending implicatures and new ability

Table 1 Christine's C-DA scores

Pretest score (12)	Near-transfer posttest score (12)	Far-transfer posttest score (12)
7	11	11

Note. The maximum test scores on each part of the test are provided in parentheses

emerged through this process. Interestingly, she was able to produce the same high score (11 out of 12) on the far-transfer posttest as the near-transfer posttest. As van Compernelle and Zhang (2014) claimed in their case study, each successive task represents a context for the transfer of capacities that were mediated in previous tasks. In view of the fact that the far-transfer task in the current C-DA involved challenging indirect opinion items, Christine’s test scores show that she was able to transfer the newly-formed ability to more complex contexts, additional evidence of her development during C-DA.

4.2 *Microgenetic Development Mediated by C-DA*

Along with the test scores, Christine’s response processes were analyzed for insights into her microgenetic development (i.e., an individual’s moment-to-moment changes in functioning) mediated by C-DA. The analysis followed a multimodal discourse analysis framework, a paradigm in discourse studies that combines language and other resources (e.g., images, gestures, actions, etc.; O’Halloran, 2011). This approach allows researchers to trace the complexity of language learning and development from various angles. In particular, the analysis centered on Christine’s engagement with C-DA during the intervention session, the stage when instructional prompts were provided to mediate her understanding of the concept of indirectness in Chinese.

Recall that Christine answered seven out of 12 questions correctly on the pretest, which focused on indirect refusals. Therefore, in the intervention session she only encountered mediational prompts on the five questions she answered incorrectly. Excerpt 1 below displays how she interacted with C-DA when she dealt with the first item in the intervention session. The aural text for this item was “Wenzhong invites his friend, Xiaomei, to go hiking during the weekend. Xiaomei says, ‘这个周末我要复习考试 (*I will review for an exam during the weekend*).’” In the pretest, Christine’s response on this item was 92, which was very close to the threshold score for the same item, 93. Figure 2 provides an example of ScreenFlow’s interface during Christine’s interactions with the same item during the intervention session.

Excerpt 1

Line	Time stamp	Description
1	10:16—10:26	While the audio plays, Christine’s eyes move on the screen frequently
2	10:26	Audio ends
3	10:27—10:28	Christine clicks “Show the Question” A MCQ appears on the screen
4	10:28—10:29	Christine moves the cursor on the screen
5	10:29—10:31	Christine stops the cursor at an empty space on the screen
6	10:31—10:32	Christine moves the cursor to the literal meaning option, “Xiaomei will review for an exam during the weekend,” and clicks it

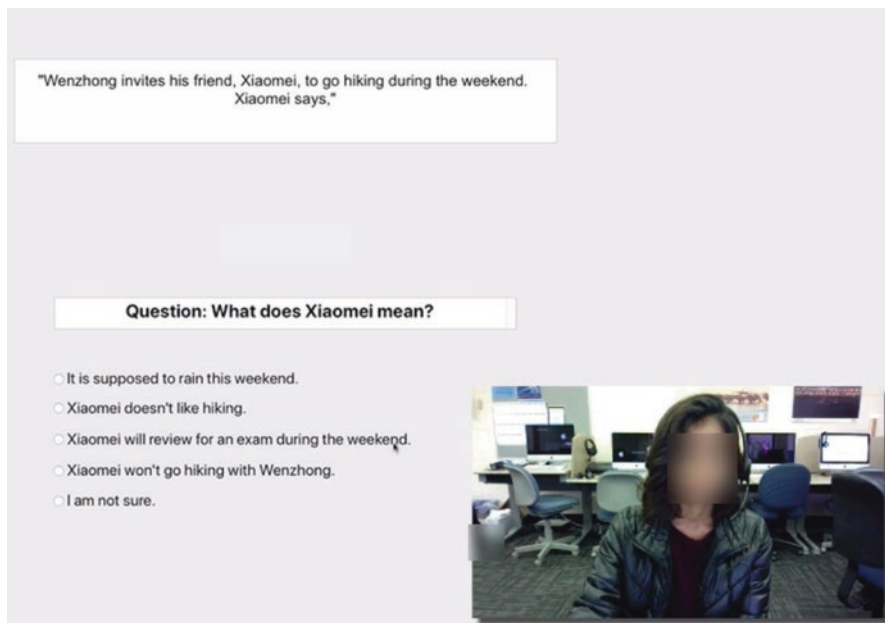


Fig. 2 ScreenFlow's interface during Christine's interactions with C-DA

Line	Time stamp	Description
7	10:33—10:36	Christine moves the cursor slowly to "Next" and then clicks it
8	10:37	The first prompt "That is not right. Try again" is shown on the screen
9	10:38—10:40	Christine's eyes widen a bit and she clicks "OK" Audio starts to play again
10	10:40—10:41	Christine's eyes move to the left quickly and then move back to look at the MCQ on the screen
11	10:41—10:44	While the audio plays the background of the conversation in English, Christine's eyes move to the right and then back to the center of the screen
12	10:45	Christine clicks the implied meaning option "Xiaomei won't go hiking with Wenzhong"
13	10:45—10:49	Audio plays the target sentence "I will review for an exam during the weekend" (<i>in Chinese</i>). At 10:47, Christine's eyes move to the left and then back to the center of the screen
14	10:49—10:50	Audio ends Christine clicks "Next"
15	10:51	The feedback "Correct!" appears on the screen
16	10:52	Christine clicks "Next"

In Excerpt 1, Christine concentrated on listening to the content and thinking (line 1) and did not look at the MCQ until the audio was done (lines 2–3). It then took her eight seconds to select the literal meaning of the utterance (lines 4–7). To

be specific, she moved the cursor on the screen (line 4), stopped there for two seconds (line 5), moved the cursor to the literal meaning option (line 6), and moved the cursor slowly to the “Next” button to check the answer (line 7). This process indicates her hesitation – it seemed that she was not very confident about the literal meaning option being correct. When the first prompt appeared on the screen, her eyes widened (line 9). We can tell that she was a bit surprised, but not too surprised. This might be due to the fact that her initial response on the same item in the pre-test was very close to the correct answer. While the audio played for the second time, she seemed to be thinking hard as her eyes moved back and forth on the screen (lines 10–11). After this short thinking/reflecting/learning process, she moved the cursor to the implied meaning option (the correct answer) before the target Chinese sentence was even played (line 12). She finished listening to the audio and then swiftly clicked the “Next” button to check her answer. Obviously, this time she was more confident about her answer.

Christine’s interactions with C-DA give insights into her microgenetic development in implicature comprehension. She only struggled a bit on the first item in the intervention session, which is largely because her initial independent response on the same item was very close to the correct answer. The mediation through C-DA helped her move forward, and one prompt was sufficient for her to develop awareness of indirectness in Chinese. Even though she did not need the second prompt that emphasizes literal utterance (i.e., the locutionary act), nor the third prompt that underscores the intention of the speaker (i.e., illocutionary act), the computer mediator still played a crucial role that supported her in linking her knowledge of indirect language use to her C-DA performance. Her performance on the remaining four items in the intervention session supports this claim. She did not need any mediational prompts on any of the four items. Particularly, on the last two items, she selected the implied meaning options before she even listened to the target sentences in Chinese. Christine’s performance on these items in fact provides additional evidence of her developing implicature comprehension ability. As Aljaafreh and Lantolf (1994) stated, changing qualities of support required from the mediator can also reveal learner development.

In addition, Christine’s response processes also reveal how she moved towards greater independence during C-DA, further evidence of her microgenetic development. Christine had a ZPD for implicature comprehension in Chinese, and the reasons are twofold. First, Christine was hesitant about choosing the literal utterance as the correct answer in the first attempt. Second, when she saw the prompt saying that her initial attempt was incorrect, it only took her a few seconds to select the implied meaning option without listening to the target sentence one more time. With that being said, her knowledge on implicatures was not fully-formed initially and thus could not be revealed in her independent performance until external support from C-DA was given. The test helped her build the connections between her existing knowledge of implicatures and the tasks at hand, which explained why she became more confident and arrived at the correct answers on the rest of the items almost effortlessly.

4.3 *Metacognitive Reflections on Implicature Comprehension*

Schmidt's (1993) noticing hypothesis states that attention is necessary for input to become noticed and then internalized. In Christine's case, the computerized mediation drew her conscious attention to the pragmatic tasks at hand, and then her metalinguistic knowledge of indirect language use came into play. Excerpt 2 displays Christine's reflections on C-DA during the interview. We can see that Christine is a reflective learner. When the researcher asked her thoughts on the functions of the intervention session, she answered, "*after that point, you would've realized that you have to answer indirectly*". Christine's initial thoughts about the test can be interpreted as her growing awareness of indirect language use. One may be concerned, though, that the thoughts may simply be a result of the test-taking strategies she developed. However, her follow-up answers definitely reveal that she started to comprehend the concept of indirectness in Chinese. Specifically, when asked whether the intervention session was helpful for her, she answered, "*yeah, cause that's when I definitely realized that I should be inferring the meaning*".

Excerpt 2

1	Researcher:	Why do you think we put the MCQs there?
2	Christine:	I think it was to, after that point, you would've realized that you have to answer indirectly because when, when you answered based on exactly what they said in the audio file, it would say incorrect, but then if you answered something else, like what you took indirectly from the audio file, it would say correct.
3	Researcher:	Do you think the MCQs were helpful for you to complete the rest of the test?
4	Christine:	Yeah, cause that's when I definitely realized that I should be inferring the meaning.

In sum, Christine's emerging implicature comprehension ability mediated by C-DA can be identified from her quantitative test scores, her moment-to-moment response processes, and her reflections upon the learning process during the exit interview. A combination of various sources of data collaboratively reveal how she started to understand indirectness and how she transferred the newly-developed ability across individual items and throughout the test.

5 **Pragmatic Instruction in Classrooms and a Boundary-Crossing Pedagogical Tool**

As the preceding discussions have made clear, the C-DA instrument helped Christine notice and pay attention to both literal and implied meanings, and she gradually understood indirect speech acts in Chinese. It is noteworthy that her marked improvement happened within a single test setting. However, Christine's quick responses and self-reflections might not be representative of all Chinese learners. In

another pragmatic intervention study reported by Qin and van Compernelle (2021), the researchers found large individual variation in terms of participants' responsiveness to computerized mediation – some participants responded well and made significant progress, while some did not respond and remained at the initial level. Is it related to their formal language learning? Do Chinese instructors intentionally emphasize appropriate language use in class? Do pragmatic concepts appear in Chinese textbooks? Apparently, these are the questions that need to be extensively researched and cannot be answered in this chapter. However, Christine's interview data revealed, perhaps not in a representative way, some current situations of pragmatic instruction in Chinese classrooms.

Specifically, Christine shared her experience of practicing a polite form in Chinese class (see Excerpt 3). She mentioned that she learned a polite way to make a request (turn 2), but “*never have conversations like this where you have to infer the meaning*” (turn 4). Moreover, she also mentioned that her teacher usually expected direct answers from students: “*our teacher says you have to be more specific about what you mean*” (turn 4). This is completely understandable owing to the fact that proficiency-oriented language teaching guidelines generally dominate the field of L2 education, and language instructors tend to focus the tight class time on students' learning of ‘essential’ linguistic patterns (e.g., vocabulary, grammatical structures).

Excerpt 3

1	Researcher:	Have you learned something related, for instance, making a request, in Chinese class? Like, can I borrow your car?
2	Christine:	Yeah, we learned very polite, direct ways to ask the question, like using ‘麻烦 (May I trouble you...)’.
3	Researcher:	Have you learned how to respond to people's request, invitation, etc.?
4	Christine:	Actually no. We never have conversations like this where you have to infer the meaning. Because if you do that in class, our teacher says you have to be more specific about what you mean. In class I think it is not accepted, but maybe outside of class it is more natural. Otherwise, sometimes they might think you don't understand what is being asked if you don't answer directly.

With regard to the extent to which pragmatic features appear in Chinese textbooks, based on my observations as a language instructor, for instance, the elementary-level textbook I used for several years (Wu et al., 2010) involves a few pragmatic features, like polite ways to make a request to borrow someone's car and invite a friend to a birthday party. Nevertheless, appropriate/polite ways to turn the requests/invitations down (e.g., using implicatures) are not mentioned. Interestingly, Christine's reflections as a language learner and my observations as a Chinese instructor align with research on L2 instructed pragmatics. Taguchi's (2015) state-of-the-art article claimed that the scope of target pragmatic features in instructed pragmatics research is skewed toward some speech acts (e.g., request, apology, etc.), and the teaching of other pragmatic targets, like implicatures, is still underexplored.

There is no argument that more weight should be given to pragmatic instructions. For instance, a wider range of pragmatic features should be taught, materials in addition to textbooks should be developed, and so on. However, given the demanding nature of language learning and limited class time, the field is calling for innovative ways to teach pragmatics to fill the gaps mentioned earlier. In response to the current situation, the C-DA instrument crossed boundaries by providing a computerized pedagogical tool to mediate Chinese learners' understanding of indirectness at any time and in any place. The computer mediator is also able to offer pre-scripted, standardized prompts to a large group of L2 Chinese learners simultaneously. Therefore, instead of using class time to explain relevant concepts, learning them through C-DA online can be very efficient and helpful for Chinese learners. In short, boundaries in learning environments and traditional teacher-student roles were crossed.

6 Conclusion and Future Directions

This chapter reports on the design and administration of a C-DA instrument that aimed to evaluate and promote L2 Chinese learners' pragmatic comprehension ability. A case analysis of an intermediate Chinese learner's emerging understanding of indirectness in Chinese was presented. Following Vygotskian genetic method, multiple sources of data (quantitative test scores, the participant's response processes during C-DA, and her reflections on the computerized test during interview) were analyzed in order to expand the evidential basis of learner development mediated by C-DA. The participant's marked improvement in a short period of time demonstrates the effectiveness and efficiency of C-DA as a pedagogical tool.

Taguchi (2015) recognized the potential of technology in expanding the options of pragmatics instruction and providing a solution to existing barriers to formal teaching. This chapter echoes her point. The current C-DA provides an efficient way for Chinese learners to develop pragmatic comprehension ability and leads to assessment innovations (Chapelle & Voss, 2016). Following Qin and van Compernelle (2021), the current computerized mediational tool has the potential to distribute assessment and instructional labor. For instance, language instructors can use this C-DA test as an assignment to examine students' independent performance and the extent to which they are able to make progress via computerized mediation. The information gathered will be helpful for instructors to plan future teaching.

However, understanding the concept of indirectness is not the most demanding task for L2 learners. What is equally important for language learners is to be able to produce indirect speech acts in their L2. Ellis (2005) argues that a student's conscious knowledge of language does not necessarily equate to ability for use. Taguchi (2016) narrows this down to pragmatics by arguing that discrepancies might exist between L2 students' knowledge regarding normative use of pragmatic conventions and how they actually implement this knowledge in real-life scenarios. Future research can explore potential ways to design computerized mediational tools to

assist L2 learners in producing implicatures. Such research is much needed and will continue crossing boundaries by exploring effective instructional tools that are able to guide students to apply indirect meanings in different social contexts.

Appendix A: Background Questionnaire

- (1) Gender _____
- (2) Age _____
- (3) Dominant language _____
- (4) Other languages: _____
- (5) What level of Chinese classes have you already completed at your university?
 - Elementary Intermediate Advanced Other (specify) _____
- (6) What level of Chinese class are you currently taking at your university?
 - Elementary Intermediate Advanced Other (specify) _____
- (7) Number of years of formal Chinese instruction _____
- (8) List the following information for any previous Mandarin studies (e.g., college, high school, intermediate/elementary school, Chinese language school, private language institute, private tutor, etc.). Please also include the current study program.
 - a. School: _____ (school name) in _____ (country name)
 Start year: _____ End year: _____ Hours of Mandarin class per week _____
 - b. School: _____ (school name) in _____ (country name)
 Start year: _____ End year: _____ Hours of Mandarin class per week _____
 - c. School: _____ (school name) in _____ (country name)
 Start year: _____ End year: _____ Hours of Mandarin class per week _____
 - d. School: _____ (school name) in _____ (country name)
 Start year: _____ End year: _____ Hours of Mandarin class per week _____
- (9) Have you visited/lived in a Chinese-speaking country?
 - No Yes (if yes, see below)
 - At what age _____ Length of stay _____ Location _____

At what age _____ Length of stay _____ Location _____

- (10) Do you speak or use **Chinese** (including **Chinese dialects**) at home? _____
- (11) Do you identify yourself as Chinese? _____
- (12) Check if your parents, grandparents, or anyone else in your immediate/extended family is a native speaker of Mandarin Chinese or a Chinese dialect.
- Mother Father Maternal grandparent(s) Paternal grandparents(s)
- Other(specify) _____

Appendix B: Interview Questions

Reflections on C-DA instrument

1. What do you think about the test you just took in general?
2. What do you think the test was evaluating? (Explanations of the test provided here)
3. Is implied meaning commonly used in your native language? While you were taking the test, did you refer to your native language experience to help you understand the implied meaning in Chinese? If so, how?
4. Is there any part of the test that was particularly challenging for you? Can you explain why?
5. The second part of the test included multiple-choice questions, and hints were provided when you didn't choose the correct answers on your own. Did you find the hints helpful? Why?
6. What did you do when you were not sure about an answer?
7. Did you use any strategies while taking the test? What kind of strategies?

Chinese learning experience

1. Can you reflect upon which aspects of your Chinese learning experience were helpful when completing this assessment?
2. Do you think taking Chinese classes helped you, or did your practice outside of class help you? Why?
3. Which aspect of Chinese is specifically challenging for you? Why?
4. Questions related to instructional level:
 - (1) How long have you learned Chinese in school (k-12, university)?
 - (2) What level of Chinese class are you currently taking?
 - (3) What activities in your Chinese class particularly contribute to your learning?
5. Questions related to heritage status:
 - (1) Do you speak or use Chinese at home? With whom?

- (2) How often do you use Chinese at home?
 - (3) Do you identify yourself as Chinese? Why?
6. Questions related to study abroad:
- (1) Have you ever visited a Chinese-speaking country? When? Where? For how long?
 - (2) Have you ever participated in a study abroad program in Chinese-speaking countries? If so, could you tell me about your study abroad experience.
 - (3) Did you have opportunities communicating with native speakers?
 - (4) Did you make friends with Chinese native speakers while studying abroad? If so, did you spend a lot of time practicing Chinese with them?
 - (5) Do you think your Chinese improved a lot after the study abroad experience? Which part improved the most?
7. Questions related to other second language learning experience:
- (1) How many languages can you speak?
 - (2) When did you learn them? For how long?
 - (3) Of all the languages you speak, what do you think of Chinese? Is it hard to learn? Why?
8. Other related questions:
- (1) Why did you learn Chinese?
 - (2) Do you have Chinese friends at your university? Do you practice Chinese with them?
 - (3) Outside of your Chinese class, do you have other opportunities to practice Chinese? How?
 - (4) How do you evaluate your Chinese skills (listening, speaking, reading, writing)?
 - (5) Is there anything else you would like to add about your experience learning Chinese?

References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78, 465–483.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in Language Learning & Technology. *Language Learning & Technology*, 20(2), 116–128.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172.
- Feuerstein, R., Rand, Y., & Hoffman, M. B. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques*. University Park Press.

- Grice, P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (pp. 41–58). Academic.
- Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational application*. Cambridge University Press.
- Haywood, H. C., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education*, 77(2), 40–63.
- Lantolf, J. P., & Poehner, M. E. (2010). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*, 15(1), 11–33.
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
- LoCastro, V. (2011). Second language pragmatics. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 319–344). Routledge.
- Nair, I., Norman, M., Tucker, G. R., & Burkert, A. (2012). The challenge of global literacy: An ideal opportunity for liberal professional education. *Liberal Education*, 98(1), 56–61.
- O'Halloran, K. L. (2011). Multimodal discourse analysis. In K. Hyland & B. Paltridge (Eds.), *The continuum companion to discourse* (pp. 120–137). Bloomsbury Academic.
- Poehner, M. E. (2007). Beyond the test: L2 dynamic assessment and the transcendence of mediated learning. *The Modern Language Journal*, 91, 323–340.
- Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting second language development*. Springer.
- Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment (C-DA). *Language Teaching Research*, 17(3), 323–342.
- Poehner, M. E., Zhang, J., & Lu, X. (2015). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing*, 32(3), 337–357.
- Qin, T. (2018). *Computerized dynamic assessment of L2 Chinese implicature comprehension*. Unpublished doctoral dissertation. Carnegie Mellon University.
- Qin, T., & van Compernelle, R. A. (2021). Computerized dynamic assessment of implicature comprehension in L2 Chinese. *Language Learning & Technology*, 25(2), 55–74.
- Qin, T., & Zhang, J. (2018). Computerized dynamic assessment and second language learning: Programmed mediation to promote future development. *Journal of Cognitive Education and Psychology*, 17(2), 198–213.
- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 21–42). Oxford University Press.
- Taguchi, N. (2005). Comprehension of implied meaning in English as a second language. *The Modern Language Journal*, 89(4), 543–562.
- Taguchi, N. (2011). Teaching pragmatics: Trends and issues. *Annual Review of Applied Linguistics*, 31, 289–310.
- Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching*, 48(1), 1–50.
- Taguchi, N. (2016). Learning speech style in Japanese study abroad: Learners' knowledge of normative use and actual use. In R. A. van Compernelle & J. McGregor (Eds.), *Authenticity, language, and interaction in second language contexts* (pp. 82–108). Multilingual Matters.
- Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. Oxford University Press.
- Taguchi, N., Li, S., & Liu, Y. (2013). Comprehension of conversational implicature in L2 Chinese. *Pragmatics & Cognition*, 21(1), 139–157.
- van Compernelle, R. A. (2015). *Interaction and second language development: A Vygotskian perspective*. John Benjamins.
- van Compernelle, R. A., & Zhang, H. (2014). Dynamic assessment of elicited imitation: A case analysis of an advanced L2 English speaker. *Language Testing*, 31(4), 395–412.
- Van Compernelle, R. A. (2014). *Sociocultural theory and L2 instructional pragmatics*. Multilingual Matters.

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Vygotsky, L. S. (1986). *Thought and language*. MIT Press.
- Wertsch, J. V. (1985). *Vygotsky and the social formation of mind*. Harvard University Press.
- Wu, S., Yu, Y., & Zhang, Y. (2010). *Chinese link: Beginning Chinese*. Pearson.
- Zhang, H., & van Compernelle, R. A. (2016). Learning potential and the dynamic assessment of L2 Chinese grammar through elicited imitation. *Language and Sociocultural Theory*, 3(1), 99–119.

Writing Development of the Case Analysis Genre: The Importance of Feedback and Negotiated Construction in the Teaching Learning Cycle



Maria Pia Gomez-Laich, Silvia Pessoa, and Ahmar Mahboob

Abstract In universities with English as the medium of instruction (EMI), boundaries between subject/content learning and language learning are crossed and the roles of content-area faculty and English for Academic Purposes faculty are bridged. In this chapter, we report on a study in which we, writing faculty/applied linguists, crossed disciplinary boundaries by working in collaboration with Information Systems (IS) professors to develop research-informed pedagogical interventions to scaffold the writing of disciplinary genres in the undergraduate IS curriculum at an American EMI university in the Middle East. Specifically, we report on how we adapted the Teaching Learning Cycle (TLC; Rothery, 1994) to scaffold the writing of the case analysis genre in two IS courses. We document the development of two students who made effective use of explicit instruction and written feedback on drafts and consulted with a writing specialist on multiple occasions in writing the case analysis. We examine the students' texts across time and contextualize the students' development by examining changes in their drafts based on the written feedback and one-on-one consultations with the writing specialist. Through these case studies, we argue for the importance of feedback and negotiated construction in the TLC. We also discuss the importance of crossing disciplinary boundaries in EMI contexts and the strategies we have used for promoting that boundary crossing.

Keywords Teaching learning cycle · Writing development · Case analysis · Explicit instruction · Written feedback · Boundary crossing

M. P. Gomez-Laich (✉) · S. Pessoa
Department of English, Carnegie Mellon University Qatar, Doha, Qatar
e-mail: mgomezla@andrew.cmu.edu; spessoa@cmu.edu

A. Mahboob
Department of Linguistics, The University of Sydney, Sydney, Australia
e-mail: ahmar.mahboob@sydney.edu.au

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_8

1 The Teaching-Learning Cycle for Scaffolding Literacy Development: The Importance of Feedback and Negotiated Construction

In universities with English as the medium of instruction (EMI), boundaries between subject/content learning and language learning are crossed and roles of subject faculty and faculty of English for academic purposes are bridged. Over the past four years, we, writing faculty/applied linguists, have engaged in such boundary crossing by collaborating with Information Systems (IS) professors to develop research-informed pedagogical interventions to scaffold the writing of disciplinary genres in the undergraduate IS curriculum at an English-medium campus of an American university in the Middle East, where most of the students use English as an additional language. Our interdisciplinary work employs design-based research (Anderson & Shattuck, 2012), an iterative approach that involves interviewing disciplinary instructors to understand their assignments and expectations, designing and implementing an intervention (e.g., by collaboratively designing scaffolding materials), analyzing its impact on learning and teaching, and reflecting upon the intervention to improve outcomes, teaching, and theory building. Ultimately, analyzing and reflecting on the effects of such interventions can lead to “profound changes in practices” (Akkerman & Bakker, 2011, p. 146) and even to the creation of a new “boundary practice” (Akkerman & Bakker, 2011, p. 146). In this paper, we report on a study in which we worked in collaboration with IS faculty and adapted the Teaching Learning Cycle (TLC) (Rothery, 1996) to scaffold the writing of the case analysis genre in two IS courses.

The case analysis genre, a prominent genre in IS, has been defined as a “written case response in which writers analyze a case and identify key factors influencing events and actions in the case or influencing possible recommendations and decision-making” (Nathan, 2013, p. 59). In a case analysis, the student applies disciplinary concepts, theory, and knowledge to the analysis of a business/organization to identify problems or opportunities to improve and provide recommendations. Writing case analyses poses several difficulties for students. One major challenge students face is crossing the boundary from reporting on the case or summarizing disciplinary knowledge—engaging in *knowledge display*—to using the disciplinary knowledge as a lens to identify problems in the case—engaging in *knowledge transformation* (Scardamalia & Bereiter, 1987). Another challenge that students may face when writing a case analysis is the lack of explicit guidelines about the case analysis genre and its expected stages. Miller and Pessoa (2016) found that many assignment guidelines do not make explicit the stages of the case analysis and consist of a series of questions for students to answer about the case. Thus, the case analysis often looks like a question-and-answer assignment rather than a full-fledged problem-solution-analysis genre. This can be attributed to the fact that many disciplinary faculty lack explicit knowledge of features of IS genres necessary to help students understand the various rhetorical moves that are expected within their specific discourse community.

To scaffold the writing of the case analysis, we crossed disciplinary boundaries, collaborated with the IS professors, and conducted one writing workshop in two IS courses (i.e., one writing workshop per course). In our workshops, we implemented an adapted version of the TLC. Framed within Systemic Functional Linguistics (SFL)¹ genre pedagogy, the TLC is an interactive and iterative writing-focused pedagogic cycle of teaching and learning activities that includes three main stages: deconstruction, joint construction, and independent construction of text. In the deconstruction phase, the teacher engages students in analyzing a mentor text's purpose, stages, and language. In joint construction, students practice writing the target genre with their teacher in preparation for independent construction (Miller & Pessoa, 2016). Research shows that these phases of scaffolding can lead to positive writing outcomes (Humphrey & MacNaught, 2016; Mitchell & Pessoa, 2017).

The TLC's potential to enhance writing development is in part attributed to its flexibility and adaptability to the needs of students across a variety of contexts. The process is recursive and, depending on students' mastery of the genre, instruction can begin at any stage of the cycle (Rothery, 1996). As Drury (2004) argues, joint construction can be done at any point of the TLC depending on students' needs, and teachers can move back and forth between phases as appropriate. During independent construction, teachers can meet with students individually in person or online and continue to offer support and feedback (de Oliveira & Smith, 2019). Some researchers have even added stages to the TLC or, in the words of Akkerman and Bakker (2011), have created a "boundary practice" (p. 146) by adapting the TLC to their needs. For example, *collaborative construction* is a bridge between joint and independent construction in which students brainstorm and negotiate ideas with other students while teachers continue to offer support as needed (Brisk, 2014; de Oliveira & Smith, 2019).

The inclusion of further support through stages of drafting and feedback during independent construction in the TLC has been a subject of interest to SFL scholars. Drury (2004) includes both group and individual feedback in her implementation of the TLC in university biology classes. Similarly, So (2005) and Feez (2002) include peer and teacher feedback as the next step after independent construction. Most recently, the SLATE project (see Dreyfus et al., 2016), in which SFL researchers collaborated with tutors to scaffold student writing development at the City University of Hong Kong, included a *consultative cycle* within the independent construction stage. In the consultative cycle, teachers provided further support and feedback to the students. This consultative cycle was conceptualized as cycles of support through drafting, commenting/feedback, and redrafting, a process that was referred to as *negotiated construction* (Mahboob, 2014, Mahboob & Devrim, 2013). During the consultative cycle, feedback plays a very important role in helping students meet genre expectations. Thus, Mahboob and Devrim (2013) argue that

¹SFL is a theory of language developed by Halliday (1985) that focuses on the analysis of language as a meaning-making resource to accomplish different functions in different social contexts. SFL-based genre instruction aims to make language choices explicit to students and scaffold the production of increasingly complex genres.

“feedback has to play a role within the Teaching Learning Cycle” and that “it has to scaffold and support students to achieve the potential that they may not be able to achieve independently” (p. 107).

Despite the importance of feedback in aiding language development and in helping students meet genre expectations, its role is still relatively unexplored in the literature on genre pedagogy. This contrasts with the significant body of literature on feedback in Second Language Acquisition (SLA) and English Language Teaching (ELT). In SLA and ELT, feedback is often conceptualized in terms of *corrective* feedback, that is, as something that “signals an error in the usage of the second language” (Kregar, 2011, p. 3). For many researchers working in the area of corrective feedback (see Heift & Rimrott, 2008; Rezaei & Derakhshan, 2011; Sheen, 2007), the focus is mostly on syntactic and other surface-level errors (e.g., spelling, vocabulary, punctuation). For example, Kregar (2011) examined the effects of different types of corrective feedback on the improvement of learner performance of three uses of the Spanish preterite and imperfect; Rezaei and Derakhshan (2011) compared the impact of direct and metalinguistic written corrective feedback on EFL learners’ grammatical knowledge; Heift and Rimrott (2008) investigated learner responses to three distinct types of corrective feedback for misspellings produced by native English-speaking learners of German; and Sheen (2007) examined the differential effect of two types of written corrective feedback (i.e., direct-only correction and direct metalinguistic correction) on the acquisition of articles by adult intermediate ESL learners (see also Kang & Han, 2015, whose meta-analysis of 21 studies examined whether written corrective feedback can help improve L2 writers’ grammatical accuracy).

While such research on feedback in ELT and SLA is quite extensive, one problem is the narrow focus of feedback studies on syntactic errors. When drawing on work on genre studies, this narrow definition of an error is unhelpful because it only looks at students’ morpho-syntactic issues as opposed to students’ use of language across a range of strata. In addition, when working with the Teaching Learning cycle, which draws heavily on the notion of scaffolding (Vygotsky, 1978), the focus is on providing support and on scaffolding students’ language development rather than just ‘correcting.’ Once again, despite the importance of feedback in aiding writing development, in genre pedagogy, there are currently very few studies that look at feedback from a genre perspective (however, see Mahboob & Devrim, 2013). Clearly, more research on students’ writing development and their ability to meet genre and assignment expectations through analysis of whole texts is needed to show the importance of feedback and negotiated construction in the TLC.

2 The Present Study

In this study, we build on Mahboob’s argument for making feedback and negotiated construction an important part of the TLC by presenting data showing student writing development of the case analysis genre from our interdisciplinary collaboration

in two different introductory IS courses at an English-medium campus of an American university in the Middle East.² In these two courses, after crossing disciplinary boundaries and collaborating with the faculty to redesign the assignments to make expectations more explicit to students, we delivered a writing workshop in which we engaged the students in the deconstruction of mentor texts and unpacked the purpose, stages, and language of the case analysis genre.

Because of time constraints, we did not engage in joint construction. Although students went straight from deconstruction to independent construction, in course 1, the writing specialist (the first author) gave students individual written feedback using a rubric that specifically addressed the expectations of the assignment and what we taught in the writing workshop. The writing specialist also met one-on-one with several students to provide further feedback. Thus, our adaptation of the TLC valued the importance of feedback and the consultative cycle.

Overall, the explicit teaching of the case analysis genre had positive outcomes compared to case analyses produced by former students who did not receive explicit instruction. However, out of the 22 students who enrolled in both classes and who participated in our workshops, 7 students performed low in both case analysis assignments, as judged by our own analysis and the grades they received from the professors (low performance was indicated by a letter grade of C or below). Interestingly, these students did not integrate the feedback they received and did not meet one-on-one with the writing specialist. Thus, these students did not make effective use of the consultative cycle of the TLC.

This contrasts sharply with students who performed high in the case analysis assignment in the two courses. These students met individually with the writing consultant, sometimes on multiple occasions. They actively engaged in the consultative process of the TLC and jointly reflected on the feedback they received with the writing specialist.

In this paper, we focus on the writing development of two high-performing students. We use these cases to provide further support for the value and importance of feedback in the form of a consultative cycle within the TLC. We examine these students' writing development contextualizing their development based on their first draft, the feedback they received, how they incorporated the feedback, and their consultations with the writing specialist. In the next section, we describe how we scaffolded the writing of the case analysis genre deconstructing a mentor text with students using the SFL-based Onion Model (Humphrey & Economou, 2015). We also used the Onion Model to provide feedback to students and to analyze student writing development.

²These two courses ran sequentially. One was a 100-level course offered in the Fall semester and the other was a 200-level course offered in the Spring semester.

3 The Onion Model for Scaffolding and Analyzing the Case Analysis Genre

To scaffold and analyze the case analysis genre, we used the Onion Model (Humphrey & Economou, 2015). The Onion Model distinguishes “four ways in which academic writing can be patterned into discourse: through description, analysis, persuasion and critique” (Humphrey & Economou, 2015, p. 37). These discourse patterns do not occur in isolation; on the contrary, they are layered and interdependent. Humphrey and Economou (2015) acknowledge that “successful persuasive and critical writing depends upon the accumulation of knowledge developed through both description and analysis” (p. 37). It is the strategic use of description and analysis in the service of persuasion and critique that allows students to meet disciplinary expectations of writing analytical arguments and critiques.

Since our previous research (Miller & Pessoa, 2016) showed that students found it challenging to shift from reporting of the case to analysis and argumentation, our workshop focused on the language of analysis and argumentation, using the Onion Model. In course 1, students analyzed a published case study about The LEGO Group using the disciplinary framework of *innovation* and evaluated the extent to which LEGO was successful in implementing innovation. In Fig. 1, we provide a sample analysis/argument stage based on the LEGO case analysis with the kinds of questions that we used with students to inductively unpack the language of argument, analysis, and description needed to effectively accomplish this stage.³

Briefly, the text starts with an explicit positive evaluation framed within the context of the disciplinary framework of innovation: *LEGO was successful in its approach to innovation* (Pessoa et al., 2019). This explicit evaluation makes the text argumentative as the main claim is stated explicitly at the beginning of the text. The text maintains its focus on positive evaluations (*increase in profits and growth of the company's customer base*) (Pessoa et al., 2019).

As an analytical text, the text uses the disciplinary framework of innovation and focuses the evaluation on selected elements of the disciplinary framework: *LEGO was successful in its approach to innovation, particularly in its use of complementary and incremental innovation* (Pessoa et al., 2019). To arrive at this, the writer engaged in analysis that involved breaking down the LEGO case to establish what activities LEGO engaged in and how those activities fit into the elements of the disciplinary framework of innovation. The analysis resulted in the selection of complementary and incremental innovation as the two main kinds of innovation that LEGO successfully engaged in. The writer then rearranged the information from the

³We used this mentor text to unpack the argumentative analytical nature of the case analysis in the second course (see Pessoa et al., 2019). In the first course, we used a mentor text from the second course. We did this so as not to provide students with a mentor text based on the same topic or case they had to write about. For the purposes of this paper, we use the mentor text based on the LEGO case to stay consistent and to enhance reader comprehension.

Read the analysis/argument section from this sample case analysis based on the LEGO company and answer the following questions.

1. What is the writer's attitude toward LEGO's approach to innovation and how do you see that reflected in the use of language? How does this relate to the main purpose of the case analysis?
2. What are the stages of this text? How is the text within this stage organized? How is the information presented? What is the focus of each paragraph? How do the two paragraphs relate to each other?
3. What language reveals that the writer is analyzing and not describing? What is the difference between how the information is presented in the source text (the LEGO case) and how the information is presented in the case analysis?
4. What voices does the writer bring in to the text? What language reveals the integration of multiple voices into the text? How are these voices used for the main purpose of the text?
5. Do you see any places where the language reveals the writer anticipates potential disagreement from a reader?

The case of LEGO can be analyzed through the lens of innovation. LEGO was successful in its approach to innovation, particularly in its use of complementary and incremental innovation. Complementary innovation is the process of creating new products that 'complement' a company's existing products in order to enhance the original product. LEGO's use of complementary innovation was successful because it led to an increase in profits and to the growth of the company's customer base. LEGO increased its profit through complementary innovation strategies such as licensing agreements with other products, movie productions, and the opening of theme parks. LEGO obtained licensing agreements to complement its main products (the bricks) with Star Wars and Harry Potter characters and vehicles. Since obtaining this licensing arrangement, LEGO has sold over 200 million Star Wars LEGO boxes (McNally, 2016). Although some of the early complementary products that LEGO produced did not sell well (i.e., Znap), the majority of LEGO's later complementary products were well-received by the public. Also, the production of the LEGO movie and the LEGO Star Wars movie were profitable for LEGO. Lego produced its first movie in [...] In addition, the opening of LEGO theme parks around the world has contributed to LEGO's financial success. Theme Park Magazine has ranked LEGO as one of the most visited theme parks around the world. It has even gained popularity in Orlando, Florida, despite the great competition from Disneyworld and its associated parks. This shows that LEGO was successful in the use of complementary innovation as it increased the company's profits.

LEGO's use of complementary innovation was also successful because it led to an increase in the number of customers. As stated in the LEGO case, until 2011, boys made up 90% of LEGO consumers. In an attempt to broaden its appeal to more girls, LEGO launched LEGO Friends, a complementary line designed specifically for girls. The line features five core female characters and the sets include bricks that can be used to build a veterinary clinic, a beauty salon, a design studio, an inventor's workshop, etc. According to the LEGO Group chief executive officer Knudstorp, launching a line designed specifically for girls allowed the company to "succeed in reaching the other 50 percent of the world's children" (2012). This confirms LEGO's success in its implementation of complementary innovation as it helped the company to increase the number of customers.

LEGO was also successful in its use of incremental innovation. Incremental innovation is...

Fig. 1 Mentor text to scaffold the analysis/argument stage of the case analysis genre

LEGO case (based on the two source texts) for the analytical and argumentative purposes of the text.

The elements of the disciplinary framework create a taxonomy for the presentation and organization of the ideas. The text in Fig. 1 focuses on complementary innovation in two paragraphs and proceeds to incremental innovation in subsequent paragraphs. It is the fronting of the elements of the taxonomy and their unpacking

into more descriptive language with evidence to support the claims and reasons that makes this text analytical.

The text also adopts a *claim-reasons* framework within the analysis of each kind of innovation adopted by LEGO. For example, the sentence *LEGO's use of complementary innovation was successful because it led to an increase in profits and to the growth of the company's customer base* presents a claim (*LEGO's use of complementary innovation was successful*) as well as the reasons that support this claim (i.e., *increase in profits* and *growth of the company's customer base*). The elements of this claim-reasons framework are stated as abstract nouns and are linked by logical relations using *because*. Figure 2 shows a visual representation of the disciplinary and claim-reasons frameworks used in the LEGO case analysis/argument mentor text from Fig. 1. Visualizing the text's organizational taxonomy brings to light the analytical argumentative nature of this stage of the case analysis genre, which differentiates it from a descriptive text.

In our workshops, we used ENGAGEMENT, one of three main resources of the SFL system of APPRAISAL (Martin & White, 2005) so that students could understand how multiple voices are used to acknowledge source texts, show how evidence from source texts relates to their claims, and to consistently position the reader.

The writer includes voices from the readings assigned as part of the course (not explicitly stated in the text) to provide a definition of one of the elements of the disciplinary framework of innovation (complementary innovation). To support the claims being made, the writer uses evidence from the case and other sources, and

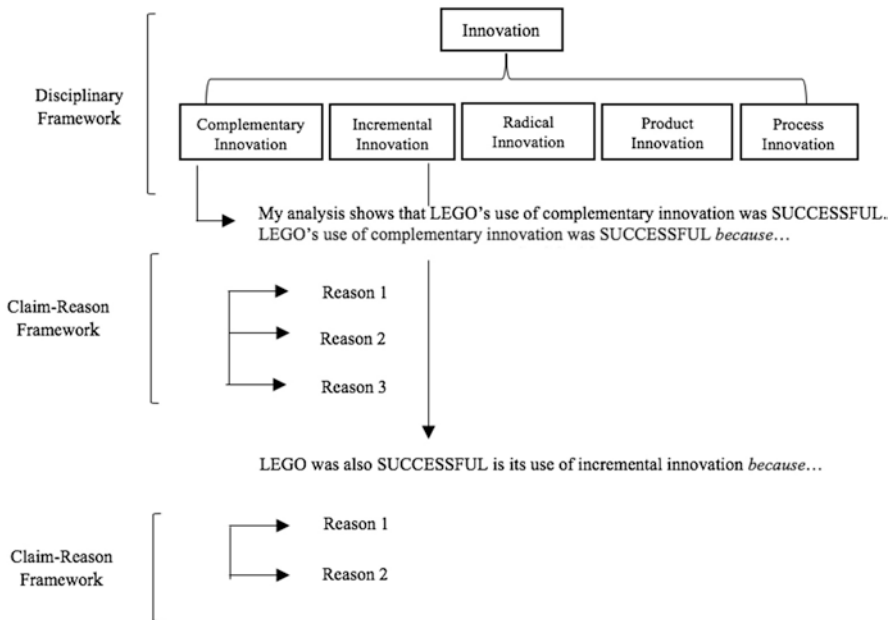


Fig. 2 Taxonomies in analytical section of a case analysis sample

signals the use of evidence with Attribute moves (i.e., moves to acknowledge the source text while introducing evidence from it) from the ENGAGEMENT system, such as: *Theme Park Magazine has ranked, As stated in the LEGO case, and According to the LEGO Group chief executive officer Knudstorp.* The student writer moves the reader toward their understanding of the LEGO case through the use of Endorsement moves (i.e., moves to show how the evidence presented supports the claim being made): *This shows that LEGO was successful in the use of complementary innovation as it increased the company's profits* (Pessoa et al., 2019). By using the Endorsement move of “*This shows that...*” the writer moves the reader toward their interpretation of the evidence presented.

In order to produce effective analysis and argument, the writer needs to (1) imagine that the reader might disagree with their analysis or evaluation, and (2) provide additional evidence to counter the alternative interpretation through the use of phrases such as *although this...that, while this...that, might*. In the text in Fig. 1, the writer acknowledges alternative perspectives and counters them in the following sentences: *Although some of the early complementary products that LEGO produced did not sell well (i.e., Znap), the majority of LEGO's later complementary products were well-received by the public, and It [Legoland in Orlando] has even gained popularity in Orlando, Florida, despite the great competition from Disneyworld and its associated parks.*

In course 2, students were not required to write a full-fledged text for their case analysis assignment. Instead, students were given a case study about SmoothPay and were to respond to a set of nine questions with thorough answers that showed the students' understanding of the text, their ability to apply disciplinary knowledge, and their research skills. In our workshop, we went through these questions and, using mentor texts, we unpacked the language for providing analytical and argumentative responses, as described earlier for course 1.

Figure 3 summarizes the linguistic features of argumentative analytical writing that we made explicit in our writing workshops and that we used in our analysis of student writing.

4 Documenting the Writing Development of the Case Analysis Genre of Two High Performing Students

In this section, we examine the writing development of two students who performed high in both case analyses. We examine how these high-achieving students incorporated the linguistic features targeted in our explicit instruction of the case analysis genre using the Onion Model. We contextualize their development based on their first draft, the feedback they received from the writing specialist, how they incorporated the feedback, and their consultations with the writing specialist.⁴

⁴This study has been approved by our institution's Institutional Review Board.

1. Clearly state an explicit evaluation (your claim) at the beginning of the text that stays consistent throughout the text.
2. If relying on a disciplinary framework, frame it as a taxonomy with specific elements, define those elements, and use the taxonomy to present and organize your ideas by fronting them in key parts of the text (usually in the beginning of paragraphs).
3. Create a claim-reasons framework and use abstract nouns to present reasons for your claim.
4. Ground your analysis and relate it to disciplinary concepts, authors, and/or incorporate evidence to support your claim using phrases such as *According to, Author X argues; X states that*.
5. Use description usually from the case strategically to support your claims.
6. Show how descriptions presented or evidence presented supports claims being made with phrases such as *This shows that, this confirms*.
7. Anticipate or acknowledge alternative perspectives and counter them with more evidence (using phrases such as *even, might, seem, although this...that, while this...that*).

Fig. 3 Linguistic features of argumentative analytical discourse in the case analysis genre

Table 1 Background information and data sources of focal students

Student	L1	High school	Data sources	Consultations with writing specialist
Samya (All names are pseudonyms)	Arabic	Arabic-medium	Course 1: Draft 1, Field notes from meetings with the writing specialist, Draft 2 Course 2: Draft, Field notes from meetings with the writing specialist	4
Yousra	Arabic	English-medium	Course 1: Draft 1, Field notes from meetings with the writing specialist Course 2: Draft, Field notes from meetings with the writing specialist	3

Before course 1, neither of the students had experience with the case analysis genre, so the explicit guidelines and their investment in negotiated construction with the writing specialist seemed to have been beneficial to these students. They both developed a working relationship with the writing specialist (the first author) and constantly sought her feedback. We use these case studies of students who made active use of the resources available to them to successfully write their case analyses to provide further support for the value and importance of feedback and negotiated construction within the consultative cycle of the TLC.

Table 1 provides background information about each individual focal student and the data sources we draw on to describe their trajectory. The two students are first year IS students from Qatar and their L1 is Arabic.

Samya attended an Arabic-medium school in Qatar and learned English in her English classes at school and at an English language school. She self-assessed her academic writing abilities, her oral presentation skills and her ability to read academic texts in English as strong (in a survey administered to all first year IS students as part of a larger study). Yusra attended an English-medium school and self-assessed her ability to read academic texts and her oral presentation skills in English as strong, and her academic writing abilities as neither weak nor strong.

Given the focus on students' ability to analyze and argue, in our analysis we focus on the analysis/argument stage of the students' case analysis in course 1, and we focus on question #6 in the case analysis assignment in course 2. Question #6 in the case analysis assignment for course 2 was the most comparable to the analytical section of the LEGO case analysis as it required students to analyze SmoothPay's competitive strategy using the disciplinary framework of Porter's Five Forces. We analyze the students' drafts and final draft qualitatively using the discourse patterns of description, analysis, and argument that we taught in the writing workshops, and that are listed in Fig. 3. We examine how the students incorporated these discourse patterns into their texts and how their writing developed as a result of the feedback received from, and the consultations with, the writing specialist.

We provide excerpts from the students' texts as appropriate, and we rely on our own field notes from the consultation sessions with these students to provide a thorough representation of these students' trajectories.

Case 1: Samya – Boundary crossing: Moving from descriptive to analytical argumentative writing through cycles of drafting, feedback, and negotiated construction with the writing specialist

Overall, Samya's first draft of the case analysis in course 1 shows that she incorporated some of the explicit guidelines we discussed in our workshop on writing case analyses, but she falls short in some aspects. (1) below shows the analysis/argument section of Samya's first draft of her LEGO case analysis in course 1.

(1) Samya's first draft of the LEGO case in course 1

When LEGO faced their first phase of the decline for the first time between 1993 and 1998, they first decided not to risk too much by pursuing incremental innovation rather than radical innovation, so they made adjustment to their existing practice by making the manufacturing process faster in order to triple the toy production, but this led to a decrease in profits.

When Plougmann became the leader and dug deep to find the problem, he chose to go with radical innovation. When Plougmann introduced the "can't miss" products, he changed the manufacturing process; the normal LEGO brick was changed into shapes which did not allow the users to create anything but the shape it was intended to be. [...] This experiment, even though it made LEGO change its core product, was a success as those sets gained huge popularity and LEGO have gained large fan bases from all over the world.

Finally, in the Galidor experiment, the basic LEGO brick was improved and an electronic system was installed in the toy which allowed users to play games. [...] This experiment was an utter failure as it destroyed a couple of actors' careers, and the show was exceptionally bad. The LEGO company had no experience in that area and did not know how to grab people's interests.

Although Samya makes an explicit evaluation and introduces the disciplinary framework in the introduction of the case analysis (not shown in (1)), she does not reiterate the explicit evaluation at the beginning of the analysis/argument section of her first draft, as we taught in the writing workshop. Overall, the analysis/argument section of Samya's first draft is quite descriptive and narrative. In (1), Samya organizes her text in terms of temporal phases (e.g., *When LEGO faced their first phase of the decline for the first time between 1993 and 1998*). Although she also refers to the disciplinary framework of innovation (notice how she mentions that LEGO first used incremental innovation and then radical innovation), it is only at the end of the first paragraph that Samya states that LEGO's use of incremental innovation *led to a decrease in profits*. In addition, Samya does not develop a claim-reasons framework to explain why LEGO's use of incremental innovation led to a decrease in profits. On the contrary, she falls into a descriptive pattern and briefly retells the activities LEGO engaged in.

In the second paragraph, Samya also uses description that is not clearly linked to the disciplinary framework or a claims-reason framework. Rather than, for example, applying an element from the disciplinary framework to show LEGO's success/failure, Samya re-tells the events of the case. Once again, it is only at the end of the second paragraph that Samya evaluates LEGO's use of radical innovation as a *success* and uses a causal link to provide a reason for why it was a success (i.e., *was a success as those sets gained huge popularity and LEGO have gained large fan bases from all over the world*). In the third paragraph, Samya loses track of the disciplinary framework, falls again into narrative and descriptive writing, and characterizes one of LEGO's experiments as '*an utter failure*'.

In the feedback given to Samya's on her first draft, the writing specialist commented that the analysis stage was too descriptive and lacked a fronted explicit evaluation. Table 2 shows the comments the writing specialist provided to the student writer.

Upon receiving the feedback, Samya met with the writing specialist twice before she submitted her final draft to the IS professor. In these meetings, they worked on making evaluations explicit, on providing relevant and sufficient evidence to support the evaluation (claim-reasons framework), and on the use of attribute moves to integrate material from outside sources into the text. Excerpt (2) below shows the analysis section of Samya's second (and final) draft of her LEGO case study in course 1.

(2) Samya's second (and final) draft of the LEGO case in course 1

As I mentioned in my introduction, my analysis shows that LEGO's use of incremental innovation in the first phase of its decline was unsuccessful because it only led to an increase in the number of products produced, and to an increase in the costs of production. This led the company to a state of financial trouble. In addition, LEGO did not succeed in its use of incremental innovation because its products did not really change and, therefore, did not adapt to the changing market. For example, [...].

In the second phase of the decline, LEGO was partly successful in its use of radical innovation because its new products, such as the sets produced in collaboration with the creators of Star Wars and the Harry Potter books, gained immense popularity. However,

Table 2 Feedback provided by the writing specialist

Comment Number	Comment
Comment 1	Samya, your analysis section is mostly descriptive and narrative. You have structured your text mostly as a sequence of events. Look at the language you are using: “ <i>When LEGO faced their first phase of the decline for the first time between 1993 and 1998,</i> ” “ <i>When Plougmann became the leader and dug deep to find the problem.</i> ” You are basically narrating everything LEGO did. For this assignment, however, you are expected to analyze and evaluate. Was LEGO successful in its implementation of innovation? When you work on your second draft, make sure you start your analysis section with an explicit evaluation of the case in relation to the concept(s) of innovation. It seems that your analysis of the case shows that LEGO was unsuccessful in its implementation of incremental innovation and was successful in its implementation of radical innovation. As we discussed in the workshop, front (place at the beginning) your evaluation (your claim) and then provide evidence to support the evaluation. So, you might want to say ‘LEGO was successful in its implementation of innovation because...’
Comment 2	Make sure you do not lose track of the concept of innovation, which is the concept that should drive your analysis. In the paragraph that starts “ <i>Finally, in the Galidor experiment</i> ” you do not refer to any form of innovation. What type of innovation does this example represent?

these products had two flaws. Firstly, they were only popular for a short period of time. Secondly, [...]. In addition, the changes in the products that LEGO introduced resulted in the loss of one of the main features of the basic LEGO brick, which is [...]. An example of such products is the computerized LEGO toys. Unlike the previous products, these products [...].

Although Samya’s final draft is still organized in terms of temporal phases, she succeeds in making an explicit evaluation at the beginning of the analysis section (*LEGO’s use of incremental innovation in the first phase of its decline was unsuccessful*) and then proceeds to give two reasons (expressed explicitly through the linguistic form *because*) of her evaluation in the form of abstract nouns (*because it only led to an increase in the number of products produced, and to an increase in the costs of production*). The second paragraph starts with a circumstantial adjunct followed by an explicit evaluation (*LEGO was partly successful in its use of radical innovation*) that makes use of the disciplinary framework of innovation. The evaluation is followed by a reason to support the claim. To address her claim that LEGO was *partially* successful in its use of radical innovation, Samya introduces a *concede* move (*However, these products had two flaws*) to provide further support for her claim. The description is now used within the disciplinary framework and in the service of the evaluation. Samya’s analysis could have been further improved if she had used a phrase such as *This shows that* or *This confirms* to show how the descriptions or evidence presented supported her claim. However, this point was not addressed in the feedback.

Samya’s case analysis in course 2 shows great improvement in her development as a writer of analytical argumentative texts. Although students in course 2 were not required to write a first draft of their SmoothPay case analysis, Samya sought the

Table 3 Feedback provided by the writing specialist

Comment Number	Comment
Comment 1	Samya, make sure you define each element of the disciplinary framework you use to evaluate SmoothPay (i.e., each of Porter's Five Forces). What does 'threat of new competitors,' 'buyer power,' etc. mean? Once you have defined the elements, state your evaluation. Is the threat of new entrants high, moderate or low? Devote one paragraph to each of Porter's Five Forces and provide evidence to support your evaluation.

assistance of the writing specialist before submitting her assignment to the professor and successfully integrated the feedback received from the writing specialist. Table 3 shows the comments provided by the writing specialist.

Samya's attention to the writing specialist's comments resulted in a case analysis for course 2 that is quite strong, as shown in (3).

(3) Samya's draft of SmoothPay case analysis in course 2

Threat of new competitors refers to the degree of likelihood that customers might switch to new companies that open in the industry. In the case of SmoothPay, the threat of new competitors is low. Even though the entry barriers are low since no substantial infrastructure is needed to start a mobile payment service, SmoothPay has loyalty programs so customers will face switching costs. Also, SmoothPay uses [...] SmoothPay could turn the threat of new market entrants to very low if they increased the entry barrier by offering even more features or differentiated services to their customers. This will increase the customers' expectation and would make it more difficult for new startups to compete with SmoothPay. An example of such features would be [...].

Intensity of existing rivalry refers to the company's position in the market in relation to its competitors. In the case of SmoothPay, the intensity of existing rivalry is high. Even though it is difficult for startups to effectively compete in the industry, the current competitors are strong and offer similar or even better services than SmoothPay. For example, [...] SmoothPay could innovate in certain ways so that users of similar services would want to switch to SmoothPay. For example, SmoothPay could combine the loyalty program, payment, and order-ahead functionality in a single transaction to gain customers that value time.

Power of suppliers refers to [...]

Threat of substitutes refers to [...]

Samya uses an effective approach to each paragraph. She starts each paragraph with a definition of each component of the disciplinary framework, then locates it in the case of Smoothpay (*In the case of SmoothPay*), and evaluates whether each force is high, medium, or low (only two paragraphs are shown as examples). In the first two paragraphs, Samya follows her evaluation with a concede-counter move (i.e., *Even though the entry barriers are low*) to back up the evaluation, and to further demonstrate that she fully understands the disciplinary framework and that her assessment is correct. She ends each paragraph providing recommendations (through the use of modal verbs) for how information systems can help the company gain a competitive advantage. Overall, Samya's case analysis for course 2 shows good evaluative patterning, good use of concrete details to support her evaluations, and good use of modals to provide recommendations. To achieve this, Samya met with the writing specialist several times. Clearly, the writing workshops, the boundary crossing

achieved by incorporating feedback in the form of negotiated construction with the writing specialist within the consultative cycle of the TLC, and her constant use of the workshop materials helped Samya meet genre expectations.

Case 2: Yousra: Boundary crossing: Moving from descriptive to analytical argumentative writing through cycles of drafting, feedback, and negotiated construction with the writing specialist

Yousra sent a draft of her assignment to the writing specialist before she submitted the first draft of the LEGO case analysis to her IS professor (course 1). In the email, Yousra explicitly stated that she was worried about using too much description and about her case analysis not being “enough analytical.” Therefore, the writing specialist mostly focused on the analysis section of the assignment. Table 4 shows comments the writing specialist made on Yousra’s draft.

Yousra attended to these comments and submitted a case analysis for course 1 that succeeds in using description in the service of analysis, as shown in (4).

(4) Yousra’s final draft of the LEGO case in course 1

LEGO’s case can be analyzed by examining two types of innovation used by LEGO to put an end to the decline in their sales. Innovation involves [...]. The products, technologies or services that result from implementing these new ideas can be completely different from previous existing versions or can involve minor improvements. In LEGO’s case, some approaches and strategies involved incremental innovation while others involved complementary innovation. Incremental innovation is [...]. In LEGO’s case, this approach was unsuccessful. LEGO abandoned their original plastic blocks to come up with toys that were designed for kids who don’t like building blocks. To illustrate this point, [...]. These toys were a complete failure because [...]. This shows that incremental innovation was not the right approach for LEGO as the new features they came up with didn’t suit their consumers.

Despite the fact that incremental innovation wasn’t successful, my analysis shows that LEGO was, overall, successful in its approaches to innovation since LEGO eventually realized what they had done wrong and they used complementary innovation to solve the issue they were facing. Complementary innovation is [...]. In LEGO’s case, this successful approach involved taking the risk to merge new technologies with LEGO’s building and creating concept [...]. This approach was successful because it had a unique aspect to it [...]

Table 4 Feedback provided by the writing specialist

Comment Number	Comment
Comment 1	Yousra, your analysis section is shaping up well. I think you can make it stronger, though. In your first paragraph, you state that LEGO was unsuccessful in its use of incremental innovation. Can you briefly state what incremental innovation is? Likewise, in the second paragraph you do not state what complementary innovation is. Remember that it is important to define the concepts that guide your analysis. You should also provide more evidence to show that LEGO was not successful in its use of incremental innovation. You state: “LEGO abandoned their original plastic blocks to come up with toys that were designed for kids who don’t like building blocks.” Can you give some examples of such toys?
Comment 2	Conclude your paragraph by showing how the evidence you provide supports the claim/evaluation you have made using phrases such as <i>This shows that...</i> , <i>This confirms that...</i>

In addition to uniqueness, this approach allows more creativity [...] Another reason this approach was successful is that recently, children are not the only LEGO fans since more adults started re-engaging with building and creating LEGO robots and structures. This benefited LEGO since [...]. Thus, this shows that the complementary innovation approach of joining technology with basic LEGO blocks has set the stage for future generation of builders, designers and innovators. Also, it helped LEGO, business wise, by strengthening the bond between LEGO and LEGO fans which in the long run increased sales.

Yousra succeeds in defining the elements of the disciplinary framework (incremental innovation and complementary innovation), in making explicit evaluations (e.g., *In LEGO's case, this approach was unsuccessful*), and in providing her own claim-reasons framework to support her asserted evaluations of LEGO's. For example, in the second paragraph, Yousra signals her overall evaluation of LEGO's success and her intention to apply one of the specific elements of the disciplinary framework (i.e., complementary innovation). She then contextualizes her analysis by accurately defining this element of the disciplinary framework according to the established knowledge of the field. Then, she shows how complementary innovation was implemented by LEGO, demonstrating that she understands this element of the disciplinary framework by applying it to information from the case.

In this stretch of text, she uses description (e.g., *LEGO used technology, LEGO transformed*) in service of her analysis that technology was the "something" that LEGO associated with its original product. Yousra moves from this analysis to supporting her asserted evaluation (*this approach was successful*). She condenses her analysis of LEGO's use of complementary innovation into a single phrase (*this approach*), and then proceeds to give three reasons for her evaluation, namely that the use of technology had unique aspects, it allowed for creativity, and was engaging to adults as well as children. Within each element of the framework she establishes, Yousra provides a claim and reason for LEGO's success (e.g., *In addition to uniqueness, this approach allows more creativity; the technology emerged with LEGO's new toys allows more space for hacking, tinkering and finding new ways of creating*), provides details from the case to support that reason, and establishes a causal link between these details and an increase in sales (e.g., *These features are great selling points, so they helped in increasing sales for LEGO*). With these causal links, she effectively uses technical language from the definition of complementary innovation to remind the reader that she is illustrating the company's successful implementation of this strategy.

Throughout this entire stretch of text, she uses explicit positive evaluations that are specifically focused on LEGO's use of technology; in other words, her evaluative position remains consistent and focused. Overall, Yousra is very effective in weaving together analytical and argumentative writing to meet genre expectations. She combines analysis using the disciplinary framework of innovation with argument using her own claims-reasons framework to support her asserted evaluation of LEGO's success/failure.

Before submitting her assignment for course 2, Yousra emailed the assignment to the writing specialist (who provided feedback) and subsequently met with the

Table 5 Feedback provided by the writing specialist

Comment Number	Comment
Comment 1	Yousra, you are making effective use of Porter’s Five Forces to analyze SmoothPay and to present and organize your information. You have fronted your evaluation of each element of the framework and you have provided evidence to support it. However, I think your analysis could be stronger if you focused more on recommendations. What can SmoothPay do to reduce industry rivalry? What can SmoothPay do to make sure the threat of new entries remains moderate/low? You can use modal verbs to do this (e.g., SmoothPay could, should...)

writing specialist. The comment provided by the writing specialist is displayed in Table 5.

Yousra attended to these comments and submitted a case analysis for course 2 that is displayed in (5).

(5) Yousra’s draft of the SmoothPay case analysis in course 2

By applying Porter’s Five Forces model on SmoothPay, it can be seen that industry rivalry has a high pressure on SmoothPay. This is because the mobile payment industry is filled with different options for customers, being merchants and consumers. SmoothPay has many competitors like [...]. All these competitors serve similar, if not the same, features and services as SmoothPay [...]. Even though not all of SmoothPay’s competitors offer loyalty rewards, most like LevelUp and Suretap do (Halliday & Dong, 2016). This makes it hard for SmoothPay to differentiate its service and products leading to fierce competition in the industry and a high pressure of industry rivalry force on SmoothPay. For SmoothPay to overcome this force, SmoothPay should widen its options by partnering with big clothes, furniture, cosmetics and sport retailers. This will encourage [...]. If SmoothPay includes other categories, customers can pay for many things at once. This will also allow more merchants to join which means SmoothPay can take advantage of the new merchants’ customer base [...]

The force of new entries is moderate to low on SmoothPay since huge capital is needed for developing, or purchasing a fully functional information system, or systems, to support the mobile payment application, or website, needed to enter this industry. Similarly, new entrants need years to establish customer trust and build a positive brand image. This is because mobile payment [...]. One-way SmoothPay can maintain the force of new entrants as low is by [...]

The buyer’s power force on SmoothPay is moderate since customers have a range of mobile payment applications to choose from. SmoothPay and its competitors offer [...]

Once again, Yousra is very effective in weaving together descriptive, analytical and argumentative writing to meet genre expectations. She combines analysis using the disciplinary framework of Porter’s Five Forces with argument (explicit evaluation) using her own claim-reasons framework to support that evaluation. She also uses modals to give recommendations at the end of each paragraph. Once again, the writing workshops and the boundary crossing achieved by incorporating feedback in the form of negotiated construction with the writing specialist within the consultative cycle of the TLC helped Yousra meet the expectations of the case analysis genre.

5 Implications for the Inclusion of Feedback and Negotiated Construction in the Teaching and Learning Cycle

In this paper, we documented the writing development of the case analysis genre among two high-performing students who made effective use of explicit instruction and written feedback on drafts and consulted with the writing specialist on multiple occasions to receive and incorporate further feedback into their texts. We examined the students' texts across time and contextualized the students' development examining changes in their drafts based on the written feedback and one-on-one consultations with the writing specialist. These case studies and their writing development provide evidence for the valuable role of feedback and negotiated construction in the TLC in helping students meet genre and assignment expectations.

Although the students' development was not always linear, these students were able to meet genre expectations due to their diligent time and effort spent on the task and their use of the resources available to them. While the former is mostly under the control of the student, the way teachers actively and systematically engage students in the writing process through cycles of drafting, feedback, and consultations is something that teachers using the TLC should strive for to enhance student writing development, as shown in this study.

Thus, like other SFL researchers (e.g., Drury, 2004; Feez, 2002; Mahboob & Devrim, 2013; So, 2005), we argue for the inclusion of feedback and negotiated construction within the independent construction cycle of the TLC. The inclusion of cycles of feedback and negotiated construction within independent construction is particularly important as explicit instruction through the deconstruction of texts does not seem to be enough for students to meet genre expectations. In higher education (where there is little time to engage in joint construction of texts), negotiated construction after the teacher has provided feedback on students' first drafts may be more feasible and beneficial than joint construction. In negotiated construction, the students are working with their own texts which is likely to generate more investment and buy-in from the students. Thus, fruitful time can be spent in the class by using examples from the students' first drafts in whole-class negotiated construction sessions, in addition to meeting with students individually if necessary.

In order to enhance this process, we understand that feedback needs to be cohesive, strategically focus on the relevant linguistic features of the target genre (Dreyfus et al., 2016; Mahboob, 2014; Mahboob & Devrim, 2013), and go beyond syntactic and other surface-level errors (e.g., spelling, vocabulary, punctuation). To make our feedback cohesive with our assignment instructions and scaffolding material, we have developed rubrics that focus on purpose, structure (stages), and development (emphasizing the use of language to develop the ideas) using the same metalanguage from our scaffolding materials. Following Mahboob's model for giving feedback, at the end of our rubric, we summarize what students are doing well in their writing and what they need to work on using the metalanguage from the rubric. Most importantly, we have argued our case with the disciplinary faculty to make time in the class schedule to engage students in negotiated construction by

working through examples of the students' first drafts. This whole-class work, in addition to individual consultations, has improved the quality of student writing. We will continue to collect and analyze student writing, our feedback, and our students' incorporation of the feedback to provide further evidence for the importance of feedback and negotiated construction in the TLC.

Finally, this study would not have been possible without engaging in a boundary-crossing interdisciplinary collaboration with the IS faculty at our institution. Despite their learning potential, such boundary-crossing interdisciplinary collaborations that bring together the English faculty's linguistic knowledge and the disciplinary faculty's content knowledge are "rare and quite demanding" (Engeström et al., 1995, p. 321) and require extensive commitment, evaluation, and continued refinement of materials and methods from all parties involved. The starting point for such boundary-crossing interdisciplinary collaborations is having an engaged and committed disciplinary faculty member interested in addressing student needs through a focus on language. Then, the language specialists need to become familiar with the demands and challenges of the professor's writing assignments, and of the discipline's linguistic and genre demands. This can be achieved by analyzing course materials and student writing (e.g., comparing high- and low-graded essays to identify valued features of the genre); conducting interviews and think-aloud protocols with the professor about student writing; and reviewing the available academic literature. These data and background knowledge form the basis for the development of the materials for an intervention. After implementing an intervention, it is important to sustain an iterative process of data collection, analysis, and reimplementation to continue refining the materials. This constant reflection might lead to "profound changes in practices" and even to the creation of a new "boundary practice" (Akkerman & Bakker, 2011, p. 146).

Acknowledgement This manuscript was made possible by NPRP grant #8-1815-5-293 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

- Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research, 81*(2), 132–169.
- Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher, 41*(1), 16–25.
- Brisk, M. E. (2014). *Engaging students in academic literacies: Genre-based pedagogy for K-5 classrooms*. Routledge.
- de Oliveira, L. C., & Smith, S. (2019). Interactions with and around texts: Writing in elementary schools. In N. Caplan & A. Johns (Eds.), *Changing practices for the L2 writing classroom: Moving beyond the five-paragraph essay* (pp. 65–88). University of Michigan Press.
- Dreyfus, S., Humphrey, S., Mahboob, A., & Martin, J. M. (2016). *Genre pedagogy in higher education: The SLATE project*. Palgrave Macmillan.

- Drury, H. (2004). Teaching academic writing on screen: A search for best practice. In L. Ravelli & R. Ellis (Eds.), *Analysing academic writing: Contextual frameworks* (pp. 233–253). Continuum.
- Engeström, Y., Engeström, R., & Kärkkäinen, M. (1995). Polycontextuality and boundary crossing in expert cognition: Learning and problem solving in complex work activities. *Learning and Instruction*, 5, 319–336.
- Feez, S. (2002). Heritage and innovation in second language education. In A. M. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 43–69). Lawrence Erlbaum Associates.
- Halliday, M. A. K. (1985). *An introduction to functional grammar* (1st ed.). Edward Arnold.
- Halliday, J., & Dong, L. (2016). *SmoothPay: Growing a mobile payment user base*. Ivey Publishing.
- Heift, T., & Rimrott, A. (2008). Learner responses to corrective feedback for spelling errors in CALL. *System*, 36, 196–213.
- Humphrey, S., & Economou, D. (2015). Peeling the onion: A textual model of critical analysis. *Journal of English for Academic Purposes*, 17, 37–50.
- Humphrey, S., & Macnaught, L. (2016). Functional language instruction and the writing growth of English language learners in the middle years. *TESOL Quarterly*, 50, 792–816.
- Kang, E., & Han, Z. (2015). The efficacy of written corrective feedback in improving L2 written accuracy: A meta-analysis. *Modern Language Journal*, 99, 1–18.
- Kregar, S. (2011). *Relative effectiveness of corrective feedback types in computer-assisted language learning*. Unpublished doctoral dissertation. Florida State University.
- Mahboob, A. (2014). Meeting the challenges of English-medium higher education in Hong Kong. *International Review of Applied Linguistics in Language Teaching*, 52, 183–203.
- Mahboob, A., & Devrim, D. Y. (2013). Supporting independent construction online: Feedback in the SLATE project. *Linguistics and the Human Sciences*, 7, 101–123.
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- Miller, R. T., & Pessoa, S. (2016). Role and genre expectations in undergraduate case analysis in information systems. *English for Specific Purposes*, 44, 43–56.
- Mitchell, T. D., & Pessoa, S. (2017). Scaffolding the writing development of the argument genre in history: The case of two novice writers. *Journal of English for Academic Purposes*, 30, 26–37.
- Nathan, P. (2013). Academic writing in the business school: The genre of the business case report. *Journal of English for Academic Purposes*, 12, 57–68.
- Pessoa, S., Gomez-Laich, M. P., Liginlal, D., & Mitchell, T. D. (2019). Scaffolding case analysis writing: A collaboration between IS and Writing Faculty. *Journal of Information Systems Education*, 30, 42–56.
- Rezaei, S., & Derakhshan, A. (2011). Investigating recast and metalinguistic feedback in task-based grammar instruction. *Journal of Language Teaching and Research*, 2, 655–663.
- Rothery, J. (1994). *Exploring literacy in school English* (Write it right resources for literacy and learning). Metropolitan East Disadvantaged Schools Program.
- Rothery, J. (1996). Making changes: Developing an educational linguistics. In R. Hasan & G. Williams (Eds.), *Literacy in society* (pp. 86–123). Longman.
- Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Cambridge monographs and texts in applied psycholinguistics. Advances in applied psycholinguistics, Vol. 1. Disorders of first-language development; Vol. 2. Reading, writing, and language learning* (pp. 142–175). Cambridge University Press.
- Sheen, Y. (2007). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41, 255–283.
- So, B. P. C. (2005). From analysis to pedagogic applications: Using newspaper genres to write school genre. *Journal of English for Academic Purposes*, 4, 67–82.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Boundary Crossing: Integrating Visual Arts into Teaching Chinese as a Foreign Language



Yan Liu

Abstract This chapter reports on the author's effort to cross disciplinary boundaries in teaching Chinese as a foreign language (CFL). It presents a mixed-methods study that examines student perceptions about, as well as the benefits and the challenges of, integrating visual arts and online art museum visits into CFL teaching. Quantitative and qualitative data were collected from a questionnaire and semi-structured interviews. Based on the findings, the author discusses the benefits of using art-integration approaches in CFL teaching, particularly their potential in answering the Modern Language Association's call for curricular transformation in collegiate foreign language curriculum (MLA, *Foreign languages and higher education: New structures for a changed world*. Retrieved from <http://www.mla.org/freport>, 2007). The author also analyzes the challenges encountered and proposes future research directions and suggestions for future integration of visual arts in the CFL curriculum.

Keywords Teaching Chinese as a foreign language · Teaching languages through art · Visual arts · Art-integration approach · CFL curriculum

1 Introduction

The Modern Language Association (MLA) of America issued a report that called for “a broad, intellectually driven approach to teaching language and culture in higher education” (MLA, 2007, p. 1) to produce “educated speakers who have deep translanguingual and transcultural competence” (MLA, 2007, p. 2). This call aimed to replace the two-tiered curriculum that separated language instruction from content, which had long been the dominant curricular template in collegiate modern

Y. Liu (✉)

Department of Asian and Middle Eastern Studies, Duke University, Durham, NC, USA

e-mail: y1286@duke.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_9

189

languages departments throughout the second half of the twentieth century (Urlaub, 2014). To answer the call, foreign language faculty need to “cross disciplinary boundaries, incorporate the study of all kinds of material in addition to the strictly literary, and promote wide cultural understanding through research and teaching” (MLA, 2007, p. 2).

Visual art materials, such as drawing, painting, sculpture, and photography, are culturally loaded. Including such materials in the language curriculum can “transform a language class into a cross-disciplinary course which not only combines language, art, literature, and history, but also helps students take a major step toward reaching proficiency standards.” (Ortuño, 1994, p. 500).

Previous studies have shown that integrating visual art or art museum visits into language curriculum (hereafter referred to as the art-integration approach) benefits both first language (L1) learning (e.g., Brouillette, 2012; Gambrell & Koskinen, 2002; Wilhelm, 2004) and second language (L2) or foreign language (FL) learning (Berhó & Defferding, 2005; Díaz, 2016; Knapp, 2012; Ortuño, 1994; Sederberg, 2013; Spina, 2006). For example, Ortuño (1994, p. 501) outlined the following eight benefits related to teaching Spanish as an L2 through visual art: (1) providing students with a social, historical, geographical, and religious context for using their language skills; (2) lowering students’ learning anxiety; (3) broadening students’ cognition at all ability levels; (4) creating a learner-centered learning environment; (5) improving students’ analytical thinking skills; (6) increasing students’ learning interest and active exploration of problems; (7) facilitating the transition from language acquisition to literature learning; and (8) motivating students to continue learning foreign language and culture because of the sense of accomplishment when they are capable of talking about art in a meaningful way.

Notwithstanding the various benefits, the art-integration approach has not been popular among Chinese as a foreign language (CFL) instructors. Few pedagogical reports or empirical studies have been published about applying this approach in the collegiate CFL curriculum. Moreover, although there are some publications on teaching Spanish (e.g., Berhó & Defferding, 2005; Díaz, 2016; Ortuño, 1994) and German (e.g., Knapp, 2012; Sederberg, 2013) with the art-integration approach in U.S. colleges, most of them are pedagogical reports that relied on students’ informal feedback or the author’s classroom observation as the only data source. Therefore, empirical studies are needed to reveal the benefits and challenges of implementing this art-integration approach into the post-secondary CFL curriculum.

To bridge the practice and research gaps, the researcher of this study experimented with the art-integration approach in a college-level CFL course, aiming to cross boundaries as discussed above in the MLA’s calls for language course innovations. This study examines student perceptions about, as well as the benefits and the challenges of, integrating visual art and online art museum visits into CFL teaching and identifies areas that need improvement for future endeavors. Different from previous studies, this study adopted a mixed-methods research methodology. The findings of this study are expected to shed some light on the development of an interdisciplinary CFL curriculum and on crossing boundaries for innovation in language education.

2 Literature Review

Previous research has found that visual arts can motivate students' learning and increase class attendance (e.g., Baker et al., 2002; Brouillette, 2012; Catterall & Peppler, 2007), communication skills (e.g., Podiozny, 2000), critical thinking skills (e.g., Bowen et al., 2014) and creativity (e.g., Moga et al., 2000). Visual arts have also been found to be beneficial for first language development in K-12 contexts (e.g., Brouillette, 2012; Gambrell & Koskinen, 2002; Wilhelm, 2004).

The idea of teaching an L2 or FL through art is not new. Little research, however, has been published about the actual use of art in L2 acquisition (Berhó & Defferding, 2005). To date, only a handful reports have been published where this approach has been adopted to teach English (Spina, 2006), Spanish (Berhó & Defferding, 2005; Díaz, 2016; Ortuño, 1994), and German (e.g., Knapp, 2012; Sederberg, 2013) as an L2 or an FL.

Spina (2006) investigated whether an art-based ESL program would facilitate English language acquisition with a quasi-experiment. Two classes of 5th-grade Spanish-speaking students participated in this experiment. The treatment group attended two art-based classes every week, whereas the control group was taught with traditional ESL methods. Quantitative data (including students' scores in English and Spanish achievement tests and teacher questionnaires) and qualitative data (including observations, audio recordings, and teacher interviews) were collected and analyzed. Results showed that an art-based language curriculum facilitated ESL students' English learning without sacrificing their first language proficiencies.

In addition to ESL, the art-integration approach has been adopted in teaching Spanish as a foreign language. For example, Ortuño (1994) illustrated how to incorporate artworks into different levels of Spanish language courses, a cross-disciplinary course, and a study-abroad program for undergraduates. Based on her observation and teaching experience, Ortuño asserted that using paintings at any level of language instruction had various pedagogical benefits. To name a few, it allowed students to practice their oral and written skills in a "social, historical, geographical, or religious context" (p. 501). It engaged students and lowered their learning anxiety. It also created a learner-centered learning environment, developed students' analytical thinking skills, and motivated students for further language and culture study.

Berhó and Defferding (2005) reported on students' responses to art and art projects in college-level Spanish and French classes. Based on the authors' classroom observations, students enjoyed learning the target culture through art, especially those who were shy or at a lower proficiency level than the rest of the class. Moreover, students learned new vocabulary and were motivated to talk through creating an art project. Berhó and Defferding also demonstrated how to integrate art and student-created art projects into Spanish and French classes with detailed examples. They suggested that teachers should first activate students' background knowledge on the subject and then ask students to learn about the artwork and artists, or

even ask students to create their artwork and then describe it or discuss it orally or in writing.

Another publication that documented the art-integration approach in the Spanish curriculum is Díaz (2016). In this study, Díaz described how she incorporated art museum visits into a Spanish business course at a liberal arts college. Based on email communications with students, faculty, and the director of the campus gallery, the author concluded that supplementing the foreign language curriculum with the incorporation of art museum visits had benefits for all shareholders. Díaz also argued that integrating art museum visits into foreign language curriculums could offer educators the same benefits that service-learning can.

Inspired by Ortuño (1994), Knapp (2012) demonstrated how to integrate visual art in teaching undergraduate German language classes at all levels with detailed explanations of the instructional procedures, guiding questions, classroom activities, and homework assignments. Knapp suggested that paintings can be used to teach concrete vocabulary and grammatical concepts while familiarizing students with German artists in introductory courses. For the intermediate courses, Knapp suggested using paintings as a springboard for an imaginative writing project. For this project, students needed to describe the image they chose, reflect on why they chose it, draw a sketch of the artist, explain what school or movement the artist belonged to, and compare the artwork with another one either from the same period or of the same theme in another period. Knapp encouraged teachers, when teaching advanced literature or culture courses, to use visual arts to introduce aspects of literary texts, for example, using a painting or a set of paintings to introduce the key characteristics of “Romanticism” in literature.

Similarly, Sederberg (2013) showcased a fourth-semester content-based German course on modern German history incorporating museum visits in an American university. She believed that using a museum-based approach to teach a foreign language could promote students’ aesthetic and affective learning experience, provide students with opportunities to access the target culture, and connect different disciplines and contexts of learning (i.e., personal, socio-cultural, and physical spheres). Students had four museum visits, two physical visits to museums on campus and two online visits to German museums, throughout the semester. Sederberg found that the online museum visits were more challenging because students could not have the affective engagement with real objects like in an actual museum. To meet this challenge, Sederberg suggested instructors link the online museum visits to classroom discussions, homework, or projects and give students more choices in choosing online museums.

In general, all these studies reported positive aspects of integrating art or art museum visits into the L2 or FL curriculum. However, it is not yet known how to implement this approach into the post-secondary CFL curriculum and what CFL learners think of this approach. Another limitation of the previous research lies in data types and data collection. Most of the studies mentioned above tended to be proposals or pedagogical reports that relied on informal student feedback (e.g., Díaz, 2016; Sederberg, 2013) or the author’s reflections and observations (e.g., Berhó & Defferding, 2005; Knapp, 2012; Ortuño, 1994) as the only source of

information. Those efforts can barely be regarded as empirical research projects because methodological details such as data collection and analysis methods were often missing.

The current study aims to fill these research gaps. In the rest of this chapter, I will first present a college-level Chinese course that integrated visual arts and online art museum visits. Then, I will report on the evaluation of this approach in the course in light of student perceptions and the benefits and challenges of this approach based on quantitative and qualitative data collected through questionnaires and interviews. Specifically, this study focuses on the following three research questions:

RQ1. How do CFL learners perceive the approach of integrating visual arts and online museum visits into Chinese teaching in general?

RQ2. What are the benefits of integrating visual arts and online museum visits into CFL teaching?

RQ3. What are the challenges of integrating visual arts and online museum visits into CFL teaching, and what improvements can be made in the future?

3 An Art-Integrated Chinese Course

The course that integrated visual arts and online museum visits was an intermediate-level Chinese course at a private university in the southeastern US. The Chinese course was designed for non-heritage learners whose Chinese proficiencies ranged from intermediate-low to intermediate-mid based on ACTFL's proficiency guidelines (ACTFL, 2012). I, as the researcher of this study, was also the instructor of this course. The textbook used is *Integrated Chinese* (4th edition, Volume 4) (Liu et al., 2019). All lessons were covered in this course, except for the last one. The class met four times a week, 50 min on Mondays and Wednesdays and 75 min on Tuesdays and Thursdays. In the 2021 spring semester, due to the Covid-19 pandemic, all teaching at this university was conducted online via Zoom, and physical visits to the art museum on campus were impossible. However, thanks to the internet, we could cross the physical boundaries and get access to bountiful art resources online and visit art museums around the world virtually. In this semester, I designed two classes that integrated visual arts and online museum visits: one after a lesson on gender equality in China and the other after a lesson on environmental protection and energy conservation.

The first art class was 50 min long and had two art activities. The first activity centered on racial inequality in the United States through an artwork titled *Fifty Shades of White* by Jaune Quick-to-See Smith.¹ It is a map of the United States with all states in varying shades of white paint, but the neighboring countries are in multiple colors. All fifty states are all labeled in the names of white paints carried by

¹ <https://artsandculture.google.com/asset/fifty-shades-of-white-jaune-quick-to-see-smith/TQGXHvDRBFzSpw?hl=en>

hardware store brands, such as “White Peach” for Georgia and “Yucca White” for Texas. To prepare for students’ discussion about the painting, I first showed students a regular American map and asked students to introduce their hometowns to each other in small groups. After this warm-up activity, I displayed the painting and asked students to discuss it in breakout rooms of 2–3 students on Zoom. I prepared a handout that listed useful vocabulary, grammatical structures, and discussion questions to facilitate students’ discussions. Students were asked first to describe what they saw, share the names of their home states in the map of the painting, and then talk about the differences between the map in this painting and the regular American map and what the artist would like to convey through this painting. After this discussion, I showed students a third American map, which shows geographic variations of races in the United States. Races were marked in different colors. Then, students were asked to compare this map with the one in the painting and then discuss questions related to the multiracial reality and racial inequality in the United States. The second activity was an online visit to the Glasgow Women’s Library² in the UK through the online platform Google Arts and Culture.³ There were six online exhibit collections featuring feminism, gender inequality, LGBTQ, and women’s suffrage. Students were asked to first self-explore the collections, pick one favorite artwork, and then orally present it in Chinese to other students in their small groups.

The second art class was 75 min long and was focused on art and the environment. The artwork used was the collection of “Taigu Descendants⁴” by a Chinese artist named Yongliang Yang. According to the introduction on his website, he “exploits a connection between traditional art and the contemporary, implementing ancient oriental aesthetics and literati beliefs with modern language and digital techniques” (Yang Yongliang Studio, n.d.). The collection of “Taigu Descendants” has six digital collages in the tradition of Chinese landscape painting. But when looking closely, you will find that the mountains are made of skyscrapers, demolition sites, and towers, reflecting the impact of rampant urbanization and its influences on the environment. To help students appreciate his artwork, the instructor decided to let students first learn how to enjoy traditional Chinese landscape paintings. So, the class started with a warm-up activity of talking about mountains guided by questions like “Do you like mountains? Why?,” “What is your favorite mountain, and why do you like it,” “How do you feel when you see mountains?,” and “Do you know or have you been to any mountains in China?” Then the instructor briefly introduced the “Five Great Mountains of China” and the meaning of mountains in traditional Chinese culture. Following this introduction, the instructor showed two Chinese landscape paintings to students and asked them what they saw, how they felt about the paintings, and what kind of relationship between nature and humans could be inferred from the images. Students also watched a short video about how to appreciate Chinese landscape paintings. The video was in Chinese, but it had

²<https://artsandculture.google.com/partner/glasgow-womens-library>

³<https://artsandculture.google.com>

⁴<https://www.yangyongliang.com/new-gallery>

English subtitles. After watching the video, the instructor asked students to look at the images again and discuss the characteristics and cultural meanings of the traditional Chinese landscape paintings. After all these preparations, the instructor asked students to explore Yang Yongliang's artwork in small groups of 2–3, pick one they liked, compare it with the traditional Chinese landscape paintings, and then discuss what the artist would like to convey through his art. The class ended with a self-exploration activity in which students could find an artwork related to environmental protection in social media and then share it with the class.

4 Methodology

4.1 Participants

Twenty-eight undergraduate students (16 females and 12 males, with an age range of 17–21 years old) who enrolled in the aforementioned intermediate Chinese language course at a private university in the southeast of the US participated in this study. Their Chinese proficiency ranged from intermediate low to intermediate mid based on ACTFL's proficiency guidelines (ACTFL, 2012). They were from various majors, but none of them were majoring in art or art history.

4.2 Data Collection and Analysis

The present study is mixed-methods research (Riazi & Candlin, 2014). Both quantitative data and qualitative data were obtained and analyzed to answer the research questions. The quantitative data were collected from an anonymous questionnaire (see "Appendix") at the end of the semester. This questionnaire included three parts. The first part included six items asking students' perceptions about integrating visual arts and online art museum visits into Chinese learning. For example, "Do you like the approach of integrating visual art or online art museum visits into Chinese learning?" Students were asked to either give ratings on a five-point Likert scale or choose one or multiple options from a list of multiple-answer questions and provide reasons for their ratings or choices. The second and third parts included 10 questions in total, and 9 of them were open-ended questions asking for students' feedback on the two art classes and their suggestions for future improvements.

Qualitative data were collected from both the open-ended questions in the second and third parts of the questionnaire and the semi-structured in-depth interviews to enrich and triangulate the quantitative data. Five out of the twenty-eight students who completed the survey participated in an individual interview voluntarily. Two of them were female, and the other three were male. Each interview was guided by two general questions: (1) What do you think of the two classes that integrated

online museum visits and visual arts this semester? What do you like, and what do you not like? (2) What do you think of the approach of integrating visual arts and online museum visits into Chinese language courses? Based on your experience, in what aspects do you think that you have benefited from this approach? What do you think we should change or improve when using this approach in the future? The interviews were open-ended and conducted in English. The participants were encouraged to talk as much as possible about the questions and other relevant topics that they proposed. Each interview lasted 15–20 min and was audio-recorded with permission. The recordings were first transcribed and then analyzed together with the quantitative data to answer the three research questions.

5 Results

5.1 *Research Question 1: General Student Perceptions of the Art Integration Approach*

To find out how the students perceived the art integration approach, their responses to the questionnaire item “Please rate how well you like the approach of integrating visual arts or online art museum visits into Chinese learning (on a 5-point scale, 1 = not at all, and 5 = a great deal)” were analyzed. As Table 1 shows, on average, students liked the approach very much ($M = 4.46$, $SD = 0.58$). The frequency of the students’ ratings was also calculated to examine the distribution of the responses. Except for one student who rated a ‘3’, all students either liked or liked very much the approach, providing ratings of ‘4’ or ‘5’.

Based on the reasons that students provided for their ratings, students liked the approach because they thought it was “fun,” “interesting,” and “cool” in general. To be specific, they liked it because of the following three major reasons. First, they thought that it was a fun way of learning about the Chinese language and culture. Second, some students loved the “the real-life applicability” of the approach since it provided them with an opportunity to apply what they had learned in class to real life situations. A third reason was that art presented different interpretations or perspectives of critical issues like gender inequality and environmental protection.

Qualitative data from the semi-structured interview confirmed that students liked the approach because it provided them the opportunity to practice their oral skills while learning Chinese culture at the same time. For instance, one participant commented, “It is a good way for us to strengthen our speaking skills. Plus, the art-integrated classes went beyond learning Chinese vocabulary and grammar structures, and we also learned about Chinese culture.” Moreover, one participant also

Table 1 Overall student perception of the art-integration approach

Minimum	Maximum	Mean	Standard Deviation (SD)
3	5	4.46	0.58

mentioned in the interview that he liked this approach because it helped students use language in such an interesting way that it could motivate them to learn the language more.

5.2 *Research Question 2: Student Perceptions of the Benefits of the Art-Integration Approach*

Students were also asked to rate how much they thought they benefited from the art integration approach (on a 5-point scale, 1 = not at all, and 5 = a great deal). A summary of the ratings is presented in Table 2. The results show that students, in general, thought that they had benefited a lot from the approach. The frequency of students' ratings confirmed this finding. Even though eight students provided a rating of 3, the other twenty students rated it as a 4 or 5.

Another question on the questionnaire was, "In what aspects do you think that you have benefited from this approach?" Students were asked to choose all that were applicable to them from the options listed in Table 3. The percentage of each option was calculated to find out the aspects that most people found beneficial. As shown in Table 3, more than two-thirds of students thought that they had gains in their knowledge about Chinese art or art in general, motivation for learning Chinese language and culture, speaking ability in real-life situations, ability to appreciate artwork, or interest in learning Chinese arts or arts in general. However, fewer than half of the students thought they were more confident in using Chinese in real-life

Table 2 Student perceptions of how beneficial the art-integration approach is

Minimum	Maximum	Mean	Standard deviation
3	5	3.96	0.79

Table 3 Perceived benefits of the art-integration approach

Aspects	Percentages (%)
It expanded my Chinese vocabulary.	79.3
It improved my Chinese speaking ability in real-life situations.	69
It motivated me to learn more about the Chinese language and culture.	82.8
It increased my knowledge of Chinese arts or arts in general.	93.1
It developed my ability to appreciate artworks.	62.1
It raised my interest in learning Chinese art and arts in general.	62.1
It improved my confidence in using Chinese in real-life situations.	44.8
Other, please explain.	0

situations. One of the possible reasons is that talking about art in a foreign language is challenging, and many students may find it difficult. Thus, they may feel that doing art activities in Chinese cannot improve their confidence in using Chinese.

Interview data further revealed that students thought that they had benefited from the art-integration approach in many aspects. The first significant benefit was related to language learning. Students believed that the art-integration approach provided them with an excellent opportunity to practice what they had learned in new contexts. It also improved their ability to use language in unfamiliar situations, such as talking about art. For example, one participant stated, "I think my language skills were better because I was trying to use our language to describe something in-depth that's not in a textbook...we have to pull all our knowledge together." Another participant further pointed out, "It's more hands-on to talk about impactful topics through art." These quotes highlighted students' appreciation of being able to use their Chinese to talk about issues related to their life, which is essential for foreign language learners since they have limited opportunities to apply their Chinese skills outside the classroom.

The second benefit mentioned in the interview is increased knowledge and interest in Chinese art. For instance, one participant said, "One of the important benefits of this approach, for me, is that I can learn about Chinese art because I knew nothing about it before. But now, I feel like I have good background knowledge of it." Students also revealed that this approach had raised their interest in learning about Chinese art and culture because they enjoyed the Chinese artworks shown to them. As a result, some students started "looking for interesting Chinese artists and following them on social media," which they admitted that they would not have done if they hadn't gotten to know Chinese art and artists in the art-integrated classes. Several students even mentioned that they would like to visit China and see Chinese art in person after the COVID-19 pandemic.

In the interview, participants also mentioned that the art-integration approach did motivate them to learn the Chinese language and culture, but to different extents. Two participants shared enthusiastically that this approach had inspired them to further study the Chinese language and culture. They also mentioned that they incorporated what they had acquired from the art-integrated classes into their final course projects. However, the other three participants were less confirmative about the motivational benefit of this approach because they thought that not everyone enjoyed art in general or the artwork shown to them. For instance, one participant disclosed, "It does motivate me to look at Chinese art more. But I don't know if it would for most people." Echoing this statement, another participant further pointed out, "It depends on whether the artwork resonates with the student. In my case, it motivated me to learn Chinese and Chinese culture to some extent. But for other students, art may not be that interesting."

5.3 Research Question 3: Challenges of Integrating Art into CFL Instruction and Future Improvements

Despite the general positive student perceptions of the art-integration approach, three significant challenges surfaced from the qualitative data obtained from the open-ended questions in the questionnaire and the semi-structured interviews. The first challenge stemmed from the difficulties of interpreting artworks in a foreign language. These difficulties may result from students' inadequate knowledge about the artwork or of the language to talk about it. For example, one student commented, "I enjoyed the class, but I think it was complicated to talk about the artwork because I was less familiar with maps and libraries, but it was a great opportunity to apply our language to a real-life context." Another student echoed, "I think the class was very engaging, but it was challenging to use Chinese vocabulary to explain my thoughts on the artwork."

The second challenge was related to artwork selection. Ideally, the chosen artwork should be, first of all, relevant to what students are learning so that they can apply what they have learned in a real-life context. The artwork should also be interesting enough to arouse students' interest in talking or learning about Chinese. In reality, however, it is very challenging to select such an artwork since it requires teachers to put a lot of thought and time into researching what is available and then weighing the appropriateness of students' language ability and the artwork's theme before making a decision. More importantly, students in this study showed different preferences about what kind of artwork should be used in class. One student commented, "Discussing Chinese artwork is interesting since it is an area that I am less familiar with and enjoy learning more. But using Chinese to discuss non-Chinese artwork is also valuable." Another student also pointed out that, "I don't care whether we discuss Chinese artwork or western artwork. Any artwork that is relevant to what we are learning about in class would be good." However, some students cared more about learning Chinese culture than the thematic connection of the artwork to the lesson, which can be seen from this comment: "Personally, I liked learning about the artwork from a cultural standpoint, and didn't care too much about tying in course themes." As for the types of Chinese artwork, students also had different preferences. Some students preferred traditional Chinese artwork because of its rich embodiment of culture. Other students favored modern artwork (such as digital art) because it is "cool," "engaging," or "different than what we usually see or know."

The third challenge pertained to discussing sensitive topics like race and gender inequality in class. Even though students thought that the artworks used were appropriate and engaging, they felt it was hard to talk about race and gender inequality in class. As a student disclosed in the interview, one possible reason was that, "Students are always a little bit uncomfortable, a little on edge, when the issue of race comes up because firstly people don't want to say wrong things. Therefore, they are less volunteering to contribute to the discussion."

Nevertheless, all five interviewed students agreed that it was vital to talk about social justice issues, and artwork was a good discussion starter. For example, one student said, “Whatever the artists are trying to show in their artwork is a good conversation starter. That can lead to more difficult topics to talk about, like racism and things like that.” Another student echoed, “using art to start conversations about social justice issues is a very creative way, which I feel like starting to talk about is the hardest part, especially in a language class because no one knows what to say.” A third student added, “Having people talk about difficult topics through the art is a good way to make it not be judgmental or make anyone feel bad because you show a piece of art and ask people to tell their opinions on it.”

When asked about what could be improved in the future, students suggested spending more time learning new vocabulary or grammar structures before class in the future. However, they thought that the handout prepared by the instructor was helpful ($M = 4.22$, $SD = 0.94$, ratings on a 5-point scale, 1 = not at all, and 5 = a great deal). In addition, some students thought it would be good if they could learn the cultural background of the artwork before class, either by asking students “to watch a section of a movie or TV that is culturally relevant” or “to prepare materials before class (such as, find an artwork on social media before class)” or “to read a short article or watch a video that they could then reference during the discussion.” In so doing, they thought it would “make students better prepared for class discussion” and “spare more time for discussion in class.”

Moreover, students suggested, when handling difficult topics like social justice issues, teachers may “create a safe and open space for students to share their thoughts from the first day of the course,” “ask students to do some readings about the issues beforehand so that they’ll have more to bring into the discussion,” “check students’ state of mind and feelings during the discussion,” “make sure that students have learned the terminology about how to talk about the hard issues in Chinese,” and “put students into small groups of 2 or 3 for such discussions so that they feel less judgmental and more comfortable to share.”

6 Discussion

Consistent with previous studies on integrating visual arts or art museum visits into the post-secondary foreign language curriculum (Berhó & Defferding, 2005; Díaz, 2016; Knapp, 2012; Ortuño, 1994; Sederberg, 2013), this mixed-methods study found that the art-integration approach was, in general, well-received among CFL students who were learning an intermediate Chinese course at a U.S. private university. In particular, these students enjoyed learning the Chinese language and culture through art. They thought that the art-integration approach was engaging and rewarding. The visual art materials and the online museum visits spurred their interest in learning the Chinese language and culture. As Li and Zhang (2016) point out, the culture learning in current collegiate CFL curriculums focus on cultural products and practices rather than perspectives. They also reveal that the culture learning

at the elementary level and part of the intermediate level is through reading some cultural information attached to the main text. In contrast, the culture teaching at the advanced level is through learning Chinese-specific topics, such as Chinese opera. The art-integration approach piloted in this study provides students with an opportunity to learn different cultural perspectives on an issue (such as gender inequality or environmental protection in China and the U.S.), which may be one of the possible solutions to this problem.

Students in this study also appreciated the opportunity to use their Chinese language skills in real-life applications. Through enjoying and discussing artworks, they could connect their Chinese learning with real-world issues (like racism, gender inequality, and environmental protection) and other disciplines (like art, humanities, and science). This benefit of the art-integration approach, therefore, makes boundary-crossing possible for CFL teachers. With this approach, CFL faculty can develop an interdisciplinary curriculum that is culturally rich and intellectually rigorous, which will answer the MLA's call in 2007 for a curricular transformation in collegiate FL education.

Unlike previous research, this study also examined the challenges of the art-integration approach and students' suggestions for future improvement. The challenges found in this study are noteworthy because they will provide useful insights for future pedagogical applications of the art-integration approach in the CFL curriculum. First of all, the first challenge was that learning or talking about art in a foreign language may be too hard for L2 or FL learners, which is likely to be one of the CFL teachers' common concerns about using art materials. This concern is validated because some CFL students in this study did find it hard to talk about art in Chinese. However, teachers may design some warm-up activities to prepare students for later discussion of the artwork, as the instructor did in this study. Also, teachers may follow the suggestion proposed by students in this study, namely asking students to study relevant vocabulary and grammatical structures, or read some articles or watch videos about the artworks before class. Another suggestion is that teachers can ask students (especially those in lower-level Chinese language courses) to "read and discuss in English about related cultural topics while using Chinese expressions where they could, so they would not have to be limited by their Chinese language ability and could engage in more sophisticated and intellectually challenging conversations." (Li & Zhang, 2016, p. 150).

The second challenge was that teachers may find it difficult to select an artwork that is both interesting and relevant to students' learning because students may have different interests or preferences. In this study, some students preferred Chinese artworks without caring about whether the theme or the issue depicted in the painting was relevant to what they were learning. In contrast, other students cared more about the relevance than the cultural origin of the artwork. However, this challenge should not hinder CFL teachers from trying the art-integration approach because the online platforms (like Google Arts and Culture) have a vast pool of visual art materials and online art museums for teachers to use. Besides, teachers may also ask students to explore the online art collections or museums to find artwork of their

interests, do some research about it, and then present it orally in class or write an essay about it in Chinese.

Last but not least, students in this study thought that using art was a good way to start conversations about social justice issues like race and gender inequality. However, some students may find it hard to talk about such topics in language classes because they may feel uncomfortable or fear being misunderstood or judged by their peers. To meet this challenge, as students suggested, teachers may try to create an open and safe space for students to share their thoughts without being judged, have students discuss in pairs or small groups, and check with students about their feelings during the discussion. Another suggestion is that teachers may ask students to clarify or repeat their ideas in English if students think there is a chance of being misunderstood when they first expressed in Chinese.

7 Conclusion

This mixed-methods study reported on the student perceptions about and the benefits and some challenges of integrating visual arts into the post-secondary CFL curriculum. The CFL students' positive perceptions of the art-integration approach indicated the pedagogical potential of this approach in teaching CFL and transforming the CFL curriculum.

The challenges identified in this study offer valuable insights for future integration of visual art into Chinese language courses, such as preparing students for the discussion of artworks with warm-up activities or various assignments before class, exploring online resources for artwork selection or online art museum visits, asking students to select artworks of their interest, making conscious adjustments when handling complex topics, and allowing students to use English to engage students in sophisticated conversations or reduce their anxiety about being misunderstood because of their limited Chinese language ability.

However, the insights generated in this study would not be possible if boundaries had not been crossed between disciplines, research methodologies, and instructional environments. First, the art-integration approach depicted in this study has connected Chinese language education with humanities, art appreciation, and critical social issues like race and gender inequality. This innovative teaching pedagogy, if adopted, will enable language educators to cross disciplinary boundaries and develop an interdisciplinary language curriculum, which answers the MLA's call in 2007. Second, the mixed-method research design combines quantitative and qualitative research paradigms, exemplifying the boundary-crossing between research methodologies. Third, the COVID-19 pandemic has driven all educators to cross the boundaries between two different teaching environments: face-to-face teaching and online teaching. With this boundary-crossing, I was able to take advantage of online art resources and the VR technology that makes it possible to visit art museums around the world virtually. Therefore, these boundary-crossings not only made this

study possible but also deepened our understanding of the art-integrated approach and the complexity of language education.

Despite its significances, this study has a number of limitations. First, this study focused on student perceptions of the benefits of the art-integration approach, with little attention to students' learning outcomes. Second, this study did not include CFL teachers' perceptions of the benefits and challenges of the art-integration approach, which could otherwise provide additional insights. Third, although the findings of this study demonstrated the pedagogical potential of the art-integration approach, more studies are needed to examine how to use this approach in all levels of Chinese language courses and content-based courses so as to establish successful practical models for CFL teachers to adopt. All these areas need to be improved or further investigated in future research.

Appendix: Questionnaire

Part I. Art-Integrated Approach

1. Do you like the approach of integrating visual arts and online art museum visits (referred to as the art-integrated approach hereafter) into Chinese learning?

1 2 3 4 5

Not at all A great deal

2. Why do you like or not like it? Please explain.
3. How much do you think that you have benefited from the art-integrated approach?

1 2 3 4 5

Not at all A great deal

4. In what aspects that you have benefited? (Please check all the answers that apply to you)

- expanded my Chinese vocabulary
- improved my Chinese speaking ability in real-life situations
- motivated me to learn more about Chinese language and culture
- increased my Knowledge of Chinese arts or arts in general
- developed my ability to appreciate art works
- increased my interests of learning Chinese art and arts in general
- increased my confidence of using Chinese in real-life situations
- other, please explain

5. Do you suggest that we should continue integrating visual arts and museum visits into this Chinese course?
- Yes
- No
- Maybe
6. For the artwork used, which do you prefer?
- I prefer Chinese artwork.
- I prefer American artwork.
- No preference, any artwork that is relevant will be fine.

Part II. First Art Class

1. Do you like this class? Why or why not?
2. Which part of this class do you like more, the discussion of the artwork or the visit to the women's library? Why?
3. Are the visual art works and activities relevant to what we are learning? Why?
4. How useful do you think of the handout?

1 2 3 4 5

Not useful at all Very useful

5. What should we improve or change if we do similar art activities in the future? Please tell us your suggestions.

Part III. Second Art Class

1. Do you like this class? Why or why not?
2. Which part of this class do you like more, the discussion of the artwork or the visit to the women's library? Why?
3. Are the visual art works and activities relevant to what we are learning? Why?
4. Which art class do you like better, the previous one or this one? Why?
5. What should we improve or change if we do similar art activities in the future? Please tell us your suggestions.

References

- American Council on the Teaching of Foreign Languages (ACTFL). (2012). *ACTFL proficiency guidelines*. Retrieved from <https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012>
- Baker, T., Bronwyn, B., & Admon, N. (2002). *Final evaluation report on the center for arts education's New York City partnerships for arts and education program*. Center for Children and Technology.
- Berhó, L. D., & Defferding, V. (2005). Communication, culture, and curiosity: Using target-culture and student-generated art in the second language classroom. *Foreign Language Annals*, 38(2), 271–276.
- Bowen, D. H., Greene, J. P., & Kisida, B. (2014). Learning to think critically: A visual art experiment. *Educational Researcher*, 42(1), 37–44.
- Brouillette, L. (2012). Supporting the language development of limited English proficient students through arts integration in the primary grades. *Arts Education Policy Review*, 113(2), 68–74.
- Catterall, J., & Peppler, K. (2007). Learning in the visual arts and the worldviews of young children. *Cambridge Journal of Education*, 37(4), 543–560.
- Díaz, M. E. (2016). Expanding the Spanish classroom: The ‘art’ in liberal arts. *Hispania*, 99(3), 436–448.
- Gambrell, L. B., & Koskinen, P. S. (2002). Imagery: A strategy for enhancing comprehension. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 305–318). Guilford.
- Knapp, T. E. (2012). Picturing German: Teaching language and literature through visual art. *Die Unterrichtspraxis/Teaching German*, 45(1), 20–27.
- Li, Y., & Zhang, Z.-S. (2016). CFL education at the college level. In J. Ruan, J. Zhang, & C. Leung (Eds.), *Chinese language education in the United States* (pp. 141–166). Springer.
- Liu, Y., Yao, D., Shi, Y., Ge, L., & Bi, N. (2019). *Integrated Chinese* (Vol. 4, 4th ed.). Cheng & Tsui.
- Modern Language Association (MLA). (2007). *Foreign languages and higher education: New structures for a changed world*. Retrieved from <http://www.mla.org/flreport>
- Moga, E., Burger, K., Hetland, L., & Winner, E. (2000). Does studying the arts engender creative thinking? Evidence for near but not far transfer. *Journal of Aesthetic Education*, 34(3/4), 91–104.
- Ortuño, M. M. (1994). Teaching language skills and cultural awareness with Spanish paintings. *Hispania*, 77(3), 500–511.
- Podiozny, A. (2000). Strengthening verbal skills through the use of classroom drama: A clear link. *Journal of Aesthetic Education*, 34(3/4), 239–276.
- Riazi, A. M., & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues and challenges. *Language Teaching*, 47(2), 135–173.
- Sederberg, K. (2013). Bringing the museum into the classroom, and the class into the museum: An approach for content-based instruction. *Die Unterrichtspraxis/Teaching German*, 46(2), 251–262.
- Spina, S. U. (2006). Worlds together... words apart: An assessment of the effectiveness of arts-based curriculum for second language learners. *Journal of Latinos and Education*, 5(2), 99–122.
- Urlaub, P. (2014). On language and content: The stakes of curricular transformation in collegiate foreign language education. In J. Swaffar & P. Urlaub (Eds.), *Transforming postsecondary foreign language teaching in the United States* (pp. 1–15). Springer.
- Wilhelm, J. D. (2004). *Reading is seeing: Learning to visualize scenes, characters, ideas, and text worlds to improve comprehension and reflective reading*. Scholastic.
- Yang Yongliang Studio. (n.d.). Biography. <https://www.yangyongliang.com/bio>

Part III

Program Innovation, Implementation, and Evaluation

Introduction and the Tuckerian Impact

This part focuses on program innovation, development, implementation, and evaluation. Language programs, according to Norris (2016), are “intentional interventions designed to address one or more clearly defined societal needs related to language learning and/or use;” and they “present a persistently evolving constellation of factors that interact to determine often life-changing outcomes for individuals” (p. 169). Although programs vary in scale, purpose, design principles, target population, and mode of delivery and experience, among many other dimensions, they are all characterized by a principled and systematic effort to generate an impact on the target individuals, communities, institutions, and society at large. To achieve their designed purposes and maximize the impact they aim to generate, programs need to be theoretically informed, carefully implemented, and rigorously evaluated.

In the history of language education, there have been shifting perspectives on and new understandings about language, learning, use, and society. The shifts are concurrent with many interests in, and debates on, “what works” in language education. Compared to the community’s collective knowledge of diverse programs, rigorous evaluation of programs to demonstrate “what works” is scarce. Norris (2016) decries that program evaluation “largely has been ignored by the mainstream of applied linguistics” (p. 169). Program evaluation entails the multiplicity of boundary crossing that underpins this volume (see Fig. 1, Chap. 1). For example, like for program development and implementation, boundaries between stake-holders such as scholars, administrators, teachers, and communities, to name just a few, need to be crossed for evaluation. Additionally, program evaluation needs to cross paradigmatic and methodological boundaries and necessitates “mixing” methods adaptively and collecting evidence longitudinally from multiple sources, which should not be restricted to outcome measures of learning (e.g., proficiency development; Part I) but should also entail analysis of program processes (e.g., classroom processes; Part II) and views and perspectives of insiders or stake-holders (Kiely & Rea-Dickins, 2005; Norris, 2016). Rigorous program evaluation may also require

cross-linguistic and cross-contextual comparisons. To many, program evaluation could be a daunting task and only for brave boundary crossers.

In this respect, Dick's scholarship and contribution is particularly noteworthy. Dick is arguably a pioneer of language program innovation and evaluation, which is probably one of the most salient lines of his work where he made distinguished contributions to language education research, policy, and practice, and where he exemplifies the importance, possibility, and benefits of boundary crossing. As Dick argues, "*we must make a societal commitment to encourage innovative language-education programs; and [...] furthermore we should make a collective professional commitment to offer our insights and our professional expertise to articulate appropriate educational goals for our children; to help to design and implement responsive pedagogical programs, and to document and to evaluate their relative efficacy*" (Tucker, 1984, p. 159). From as early as the 1960s, Dick and his collaborators (e.g., Hayes et al., 1967) began to argue that the traditional or "standard" approach to evaluating a language teaching program, based on objective, pre- and post-intervention proficiency measures, is limited, as it "neglects highly relevant events that take place during the actual training program" (p. 23). They henceforth proposed a more process-oriented approach in which "direct observation of actual training in progress" is made to examine the extent to which "certain policies, principles, and procedures that will have been demonstrated to play a role in successful language learning" are confirmed in "course design, program administration, and individual teaching performance" (p. 23).

Dick (Tucker, 1979; Tucker & Cziko, 1978), as Norris (2016) did about 40 years later, noted the importance of program evaluation and decried the dearth of it in bilingual education. Together with his collaborators, Dick himself made great effort to bridge theories of language learning (Part I), language teaching (Part II), and policy (Part IV) in program evaluation and to promote sustainable implementation of language programs. This was manifested in his numerous scholarly activities and publications (e.g., Chenoweth et al., 2006; Tucker, 1996, 2005), but perhaps crystallized and best exemplified in the two multi-year collaborative projects on foreign language program implementation and evaluation in schools (Donato & Tucker, 2010; Lambert & Tucker, 1972). Despite their different locations—one in Montreal, Canada on immersion and bilingual schooling, and the other in Pittsburgh, United States on FLES—the St. Lambert project and the Pittsburgh FLES project both were characterized by notable crossing of boundaries between diverse stake-holders of education to implement and evaluate a program and generate impacts on students, schools, and society at large. When discussing the paucity of program evaluation studies, Dick (Tucker, 1979) alerted us to a variety of barriers, which ranged from "the transience of potential researchers" to the "unwillingness or inability of funding agencies to commit funds," and to "the unfortunate tendency for administrators to regard initial results," as opposed to longitudinal evidence, "as a major criterion" for program decisions (p. 74). This led him later to sharply point out that "research should be a collaborative activity with teachers, administrators, policy makers, and researchers serving as equal partners in the enterprise and that it is important that the concerns of disparate audiences be represented equally in the

research process;” and “the optimal planning, implementation, and dissemination of research involves, of necessity, a continuing dialogue among the diverse ‘stakeholders’ in the various phases of teaching and learning” (Tucker, 2000, p. 207).

As is exemplified in his scholarship underpinning the other parts of this volume, Dick’s scholarship in program innovation and evaluation is also characterized by crossing boundaries of quantitative and qualitative approaches and methods to generate diverse types of data to track processes of program implementation and understand program effectiveness. In addition to measuring diverse outcome variables to understand students’ language proficiency, cognitive attributes, and sociolinguistic attitudes, and tracking their development or change longitudinally (e.g., Bruck et al., 1974; Tucker et al., 1996; see also Part I), those program evaluation studies conducted quantitative questionnaire surveys on or qualitative interviews with teachers, administrators, students, and parents (e.g., Bruck et al., 1976; Campbell et al., 1974, Donato et al., 1994, 1996; Donato et al., 2000; see also Donato & Tucker, 2010) as well as qualitatively probed into classroom discursive processes and student learning (e.g., Donato et al., 2000; Pessoa et al., 2007; see Part II).

Another notable boundary crossing in Dick’s contribution to program evaluation is the purposeful comparative lens. In the evaluation of the French immersion programs implemented as part of the St. Lambert project (e.g., Bruck et al., 1974; Lambert & Tucker, 1972), English-speaking, French-learning or bilingual students’ cognitive attributes, sociolinguistic attitudes, as well as English and French language skills were compared with those of monolingual English- and French-speaking children to demonstrate the effectiveness of the program and the cognitive and social consequences of language learning and bilinguality (see Part I). In the FLES project (Donato & Tucker, 2010), on the other hand, comparisons of program implementation and impact were made between Japanese and Spanish as a foreign language. The implementation and evaluation of each language program arguably has provided a model for program evaluation (for Japanese, see Donato et al., 1996, 2000; for Spanish, see Domínguez et al., 2005; Donato & Tucker, 2007; Tucker & Donato, 1998, 2003; Tucker et al., 2001). That evaluation also generated many important insights, based on diverse data from multiple stakeholders using a variety of methods as discussed earlier, into language learning (e.g., age effects and instructional effects on language learning; Part I); language teaching, classroom processes, and teacher development (e.g., language of instruction and teacher efficacy; Part II); and policy on early foreign language learning for language majority students (e.g., program articulation; Part IV). Yet, it is the comparative lens or crossing programmatic boundaries that is particularly important for some elaboration here.

In their book on the 15-year journey in the development, implementation, and evaluation of the Japanese and Spanish FLES programs (Donato & Tucker, 2010), Dick and Rick (Donato) carefully analyze, compare and contrast, and discuss the two programs in light of “a tale of two schools” which differed in their destiny: whereas the latter sustained and flourished, the former did not and ended abruptly. Among many important findings and discussions, Dick and Rick note a number of differences between the two programs in terms of program-internal and -external

factors. For example, the implementation of the Japanese program was essentially “top down,” with the choice of language made by the executive committee of the school. There was only one Japanese teacher who was in charge of developing all curriculums and teaching them across all grade levels in the school. The teacher had limited opportunity to interact with other teachers. There was no attempt to integrate the teaching with the so-called core content subjects. The program was essentially marginalized and was not integrated into “the life of the school.” There was also a lack of program continuity and articulation across elementary (K-5; Japanese) and middle grades (6-8; French or Spanish). Not all parents were ready to be connected with their children’s learning of Japanese through the teacher’s purposefully designed interactive home assignments. In contrast, the Spanish program, from the early stages of language selection and program planning, involved active participation of multiple stakeholders, including administrators, teachers of not only Spanish but also those of other subject areas, university partners, as well as the community and parents. It was implemented from the “bottom up.” Spanish was positioned to be a core component of the school curriculum and offered through the whole duration of students’ school years. There was a group of Spanish teachers who could generate a collaborative culture. Also notably, in the Spanish program, “there was a curricular transition from a skills-based thematic unit approach to instruction in grades K-5 to a content-related approach in the middle school (grades 6-8)” (p. 42). Dick and Rick conclude that for a foreign language program to be implemented successfully and in a sustainable manner, among many other conditions, program vision needs to be shared across all stakeholders rather than imposed externally by a particular party, teachers need to be supported (e.g., linkage between the language teacher and the grade-appropriate teachers), and the language curriculum needs to be integrated into school curriculum and the program integrated into the life of the school.

The aforementioned nuanced insights into program implementation, success, and sustainability would not have been obtained if the programmatic boundary had not been crossed through a comparative lens. These insights have taught us that sustainable implementation of a program needs to attend to a myriad of program-internal and -external factors and the complex interplay between factors within and across levels, and that policies on language programs should consider not only micro-level contextual factors but also macro-level sociolinguistic, socio-cultural, economical, and political factors. For “transfer” of programs, instead of importation of models, the approach is more desirably importation of “cycles of discovery” (cf. Swain, 1996), which entails “stages and processes of evaluation, theory building, generation of hypotheses, experimentation, and further evaluation” (Tucker, 1996, p. 318) (see also Part IV).

References

- Bruck, M., Lambert, W. E., & Tucker, G. R. (1974). Bilingual schooling through the elementary grades: The St. Lambert Project at grade 7. *Language Learning*, 24(2), 183–204.
- Bruck, M., Lambert, W. E., & Tucker, G. R. (1976). Alternative forms of immersion for second language teaching. *NABE*, 1(3), 33–48.
- Campbell, R. N., Taylor, D. M., & Tucker, G. R. (1974). Teachers' views of immersion-type bilingual programs: A Quebec example. *Foreign Language Annals*, 7(5), 106–110.
- Chenoweth, N. A., Jones, C. M., & Tucker, G. R. (2006). Language online: Principles of design and methods of assessment. In R. P. Donaldson & M. A. Haggstrom (Eds.), *Changing language education through CALL* (pp. 149–168). Routledge.
- Dominguez, R., Tucker, G. R., & Donato, R. (2005). Documenting curricular reform: Innovative foreign language education in elementary school. In P. Bruthiaux, D. Atkinson, W. G. Eggington, W. Grabe, & V. Ramanathan (Eds.), *Directions in applied linguistics: Essays in honor of Robert B. Kaplan* (pp. 56–71). Multilingual Matters.
- Donato, R., & Tucker, G. R. (2007). K-12 language learning and foreign language educational policy: A school-based perspective. *Modern Language Journal*, 91(2), 256–258.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.
- Donato, R., Antonek, J. L., & Tucker, G. R. (1994). A multiple perspectives analysis of a Japanese FLES program. *Foreign Language Annals*, 27(3), 365–378.
- Donato, R., Antonek, J. L., & Tucker, G. R. (1996). Monitoring and assessing a Japanese FLES program: Ambiance and achievement. *Language Learning*, 46(3), 497–528.
- Donato, R., Tucker, G. R., Wudthayagorn, J., & Igarashi, K. (2000). Converging evidence: Attitudes, achievements, and instruction in the later years of FLES. *Foreign Language Annals*, 33(4), 377–393.
- Hayes, A. S., Lambert, W. E., & Tucker, G. R. (1967). Evaluation of foreign language teaching. *Foreign Language Annals*, 1(1), 22–44.
- Kiely, R., & Rea-Dickins, P. (2005). *Program evaluation in language education*. Palgrave Macmillan.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. House Publishers.
- Norris, J. M. (2016). Language program evaluation. *Modern Language Journal*, 100(S1), 169–189.
- Pessoa, S., Hendry, H., Donato, R., Tucker, G. R., & Lee, H. (2007). Content-based instruction in the foreign language classroom: A discourse perspective. *Foreign Language Annals*, 40(1), 102–121.
- Tucker, G. R. (1979). Bilingual education: Some perplexing observations. *Educational Evaluation and Policy Analysis*, 1(5), 74–75.
- Tucker, G. R. (1984). Toward the development of a language-competent American society. *International Journal of the Sociology of Language*, 45, 153–160.
- Tucker, G. R. (1996). Some thoughts concerning innovative language education programmes. *Journal of Multilingual and Multicultural Development*, 17(2-4), 315–320.
- Tucker, G. R. (2000). Concluding thoughts: Applied linguistics at the juncture of millennia. *Annual Review of Applied Linguistics*, 20, 241–249.
- Tucker, G. R. (2005). Innovative language education programmes for heritage language students: The special case of Puerto Ricans? *International Journal of Bilingual Education and Bilingualism*, 8(2-3), 188–195.
- Tucker, G. R., & Cziko, G. A. (1978). The role of evaluation in bilingual education. In J. E. Alatis (Ed.), *Georgetown University round table on languages and linguistics 1978* (pp. 423–446). Georgetown University Press.
- Tucker, G. R., & Donato, R. (1998). Designing and implementing an innovative foreign language program: Reflections from a school district-university partnership. *Studia Anglica Posnaniensia: International review of English Studies*, 33, 451–461.

- Tucker, G. R., & Donato, R. (2003). Implementing a district-wide foreign language program: A case study of acquisition planning and curricular innovation. In D. Tannen, & J. E. Alatis (Eds.), *Georgetown University round table on languages and linguistics 2001* (pp. 178–195). Georgetown University Press.
- Tucker, G. R., Donato, R., & Antonek, J. L. (1996). Documenting growth in a Japanese FLES program. *Foreign Language Annals*, 29(4), 539–550.
- Tucker, G. R., Donato, R., & Murday, K. (2001). The genesis of a district-wide Spanish FLES program. In R. L. Cooper, E. Shohamy, & J. Walters (Eds.), *New perspectives and issues in educational language policy: In honour of Bernard Dov Spolsky* (pp. 235–259). John Benjamins.

Student-Level Variables and Academic Achievement in a Mandarin Dual Language Immersion Program



Chan Lü, Amy E. Pace, and Liu Liu

Abstract Inspired by the pioneering work of Dr. G. R. Tucker and colleagues on French Immersion in Canada, later known as the St. Lambert experiment (Lambert & Tucker, *Journal of Educational Psychology*, 65(2), 141–159, 1972), language immersion education also proliferated in the United States in the last decades. Research in the U.S. setting has consistently reported the linguistic and academic benefits of language immersion programs, and found that dual language immersion programs produced better educational outcomes for linguistically and racially diverse groups of students compared to other forms of educational programs for language minority students. However, most of the current data come from immersion programs involving linguistically-similar languages such as English and Spanish. Drawing on three consecutive years of student achievement data from a public school providing whole-school Mandarin immersion, the current study found that while, as a group, the students made significant progresses in their scores on mathematics and English Language Arts, their academic growth trajectories did not vary based on their home language backgrounds or race. These findings suggest that Mandarin immersion, despite teaching school contents through two highly contrasting languages, is capable of providing equal access to high quality education to all children. Implications for Mandarin immersion, as well as for the field of dual language immersion education, is discussed.

C. Lü (✉)

Department of Asian Languages and Literature, University of Washington, Seattle, WA, USA
e-mail: Chanlu@uw.edu

A. E. Pace

Department of Speech and Hearing Sciences, University of Washington, Seattle, WA, USA
e-mail: Amypace@uw.edu

L. Liu

College of Education, University of Washington, Seattle, WA, USA
e-mail: liuul@uw.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_10

213

Keywords Student achievement · Home language background · Race · Dual language immersion · Mandarin

1 Introduction

For educators and researchers in the field of immersion education, the St. Lambert experiment in Quebec, Canada (Lambert & Tucker, 1972) is widely recognized as a landmark study along with many other contributions from Tucker and colleagues (e.g., Bruck et al., 1971; Genesee et al., 1975). This seminal body of work has not only given birth to immersion programs in the United States (e.g., Campbell, 1984), but also inspired the prolific research on immersion education in the following decades. Almost half a century later, this corpus remains the ultimate inspiration for the first author's own work on Chinese immersion (e.g., Lü, 2019, to which Dr. Tucker kindly provided a Foreword).

Dual Language Immersion (DLI), as it is commonly referred to in the educational literature in the United States (U.S.), is rapidly expanding; Mandarin Chinese programs (henceforth, Chinese, unless otherwise mentioned) are among the fastest growing programs in the nation. Over the last 12 years, the number of registered programs has increased from 14 to 319 (Lenker & Rhodes, 2007; see also Lü, 2019). This exponential growth is the result of various factors including state-level policy making (e.g., Rubio, 2018; Watzinger-Tharp et al., 2018, 2021) as well as community and parent support (e.g., Baig, 2011; Jaumont, 2017). Though there have been multiple different nomenclatures in the scholarship, in this chapter, we use “immersion” and “dual language immersion” interchangeably as an umbrella term for additive models of bilingual schooling (e.g., Howard et al., 2018; Lü, 2020; Thomas & Collier, 2012), unless otherwise noted.

One central concern for any pedagogical innovation is its academic effectiveness. As early as 1979, in an opinion piece on bilingual education, Tucker strongly advocated for the importance of establishing “the necessary data base upon which we shall be able to draw inferences 10 and 15 years from now concerning the cumulative impact of an educational innovation on groups of children who have completed their formal schooling and entered the workforce” (Tucker, 1979, p. 75). Since then, and especially since the beginning of the twenty-first century, there has been a growing body of research utilizing “big data” from states or school districts, which we will review below. Many studies to date have examined the benefits of DLI education, particularly from the perspective of academic achievement. However, in the U.S. context, most of such studies have been based on Spanish/English DLI programs (e.g., Lindholm-Leary & Hernández, 2011; Thomas & Collier, 2002); less well known is the impact of formal DLI schooling in other language pairs (such as English and Chinese) on students' academic performance, and how student-level variables affect such performance. The abundance of research in Spanish is unsurprising, as it is one of the most “commonly taught languages” in the American context. Such a fact leads to more available data from school districts on

DLI students. Mandarin Chinese, in contrast, is still considered a “less commonly taught language,” and it is thus more difficult to acquire a large enough sample for meaningful analysis (e.g., Fortune & Ju, 2017). This chapter, therefore, attempts to cross existing language boundaries to achieve a better understanding of the educational benefits of language immersion by examining student achievement in a Chinese DLI setting.

Given that one of the key features of DLI education is to teach school subjects through English and a partner language using no less than 50% of the instructional time (e.g., Fortune & Tedick, 2003), whether students in DLI programs can perform at grade level is of important concern, since all state-mandated assessments for school subjects, such as mathematics and English Language Arts (ELA), are conducted in English. To date, researchers have been probing this question using data from states, districts, or individual schools. Although this body of work has provided crucial support for the educational benefits of DLI education, there are still many unanswered questions especially regarding DLI programs in non-English/Spanish-DLI programs.

In this chapter, we used growth curve modeling to examine student achievement data from a high-performing and diverse urban school which offers whole-school Chinese immersion. Though the dataset is comparably small, we believe it still offers useful insights into the effectiveness of Chinese DLI programs as they continue to expand. In what follows, we first briefly review the literature on academic achievement of DLI students vs. non-DLI students before considering the role of student level factors such as home language background and race/ethnicity on DLI achievement. Throughout the chapter, we discuss the importance of crossing boundaries in DLI research to understand how these student-level variables enrich the academic context and contribute to student outcomes. We conclude by interpreting our findings within the broader landscape and future directions for DLI education across languages.

2 Academic Achievement of Linguistically Diverse Students within Immersion Education

Following Tucker’s pioneering studies such as Lambert and Tucker (1972) and Lambert et al. (1973), most studies in the U.S. have found that students in immersion programs performed as well as, if not better than, their peers in non-immersion programs. Although earlier studies were criticized for not being able to control student-level variables such as family socioeconomic status (SES), student motivation, teaching methodology, and parental attitude, recent studies have been able to utilize data based on randomized school-lottery or school propensity matching to demonstrate the educational advantages of immersion education (e.g., Steele et al., 2017; Watzinger-Tharp et al., 2018). It is perhaps safe to conclude that half a century after immersion education was first implemented in North America,

stake-holders no longer have doubts about its educational efficacy but are now primarily focused on whether immersion education can also provide equity and racial and social justice (e.g., Palmer et al., 2019; Valdés, 1997) to ensure academic success for all students. Therefore, how immersion programs serve linguistically diverse populations is a key question.

Given the linguistic context of Quebec, Canada, participating children in the St. Lambert experiment were grouped relatively neatly into English-speaking children in French immersion (experiment group), English-speaking in a traditional English-medium program (English control) and French-speaking children in a traditional French-medium program (French control). In contrast, the complex linguistic profiles among U.S. residents aged five years and older (U.S. Census Bureau, 2015) do not necessarily allow such clear-cut comparisons. In addition, DLI programs have been trying to be more inclusive to serve students of different backgrounds with mixed results (e.g., Palmer, 2010). To date, research on immersion has mainly focused on the following categories of students, aside from typically developing, English-speaking students, namely those who: (1) are of low intellectual/academic ability; (2) have special education needs; (3) have poor L1 ability; (4) are of low socioeconomic status (SES), or (5) ethnically and linguistically diverse (Genesee & Fortune, 2014). Most relevant to the current study are students who are ethnically and linguistically diverse, though these two variables are typically intertwined.

A substantive body of evaluation research has shown that DLI programs support native English speakers as well as English learners (ELs) in meeting academic goals (e.g., de Jong, 2002) and that students who are ELs in DLI programs outperform their peers in transitional programs in terms of math and reading skills (e.g., Marian et al., 2013). Alvear (2019), for example, found that Spanish-dominant ELs in two-way immersion (TWI) programs in Texas earned the highest scores in English reading at Grade 5, as compared to their peers in transitional and developmental bilingual programs. Using data from a North Carolina district which used school choice lotteries for program assignment, Bibler (2021) estimated the effect of attending Spanish DLI schools on student achievement and revealed local average treatment effects of 0.04 and 0.05 standard deviations per year in math and reading, respectively, for all students. Subsample analysis found that attending a DL school increased test scores for ELs and for non-ELs by a similar magnitude.

It is important to note that in the U.S., there is typically an overlap between low SES and linguistic minority group status, yet Serafini et al.'s (2022) study on low-income, ethnically and linguistically diverse children in Florida confirmed that participation in Spanish DLI programs (one-way immersion [OWI] and TWI) significantly predicted a unique amount of variance in fifth grade standardized reading scores and marginally significantly predicted scores in math and grade point average (GPA). Additionally, the study found DLI programs may have increased the speed with which dual language learners became proficient in English, and such expedited speed was found to be strongly associated with scores indicating academic achievement. Other researchers, for example Morita-Mullaney et al. (2021), found that when comparing student scores from Spanish DLI vs. ESL programs from the same school district in Indiana, the association between program type and

academic outcomes at Grade 4 was more pronounced for ELA, but the advantage for DLI students in terms of math outcomes at Grade 5 was not necessarily different between the two groups, owing to the fact that math instruction for both groups of students was conducted in English. The researchers cautioned that the lack of instruction in the target language (in this case, math instruction in Spanish) may reduce emergent bilingual learners' "accessibility to math content in early years which potentially has cumulative and negative effects over time" (p. 56). Thus, the researchers advocated that future investigations focus on not only "delivery time" in each language in DLI programs, but also pay close attention to the actual allocation to content areas.

A focus on Spanish as the focal language of immersion may limit our understanding of linguistic distance, children's home language experiences, and other student-level variables. Linguistically, Spanish is similar to English in many aspects; both are phoneme-based languages whose orthographies are alphabetic and the two languages also share a large number of cognates. Additionally, with a few exceptions, the population examined in the studies we reviewed here are predominantly from low SES Spanish-speaking families; therefore, the findings may not readily generalize to DLI programs involving linguistically distant language pairs with students from families of diverse racial or linguistic backgrounds. Nonetheless, a burgeoning literature points to the importance of crossing language boundaries. For example, Thomas and Collier (2012), based on their five-year evaluation study of a school district in Oregon offering Spanish and Russian DLI programs, reported that these programs were powerful in closing the achievement gaps between English majority students, non-EL language minority students, and EL language minority students on state-mandated tests. Fortune and Song's (2016) study also included information comparing subgroups within the Mandarin immersion program; they did not find any differences in math or ELA scores at Grades 3 and 5 between ethnically/linguistically diverse students or students receiving special education services in Mandarin immersion and other students in the state; however, the numbers in the immersion subgroups are extremely small ($n < 10$). Based on a subsample out of the large-scale data ($n = 93$), Steele et al. (2017) found that randomization into immersion programs as well as family-school language match together played a role in students' probability of being classified as EL. That is, for students who spoke Vietnamese and applied to Spanish programs and speakers of non-Mandarin dialects who applied to Mandarin programs, through Grade 3, those "who are randomized to immersion remain more likely than their non-immersion peers to be classified as EL in a given year, but by fifth and sixth grades, their probabilities are 6 and 14 points lower, respectively, than those of native speakers of the partner language who did not win immersion slots" (p. 302S). Taken together, this body of research provides converging evidence that despite protracted English learning trajectories for language-minority students in early grades, DLI has sustained educational benefits for ELs.

Historically, there are clear trends showing opportunity gaps rooted in race (i.e., reports from the National Assessment of Educational Progress or NAEP) and thus far, available research seems to suggest that DLI programs have been gap-closing

programs. What remains to be studied further is how student level variables which have been found to affect students' educational achievement in traditional educational programs – most notably race and home language background – affect students' achievements in DLI programs.

Since Lambert and Tucker's (1972) pioneering study in Canada, researchers in the U.S. have further investigated the effects of DLI immersion on student achievement and how student-level variables affect learning trajectories. Because the majority of currently available studies come from Spanish DLI programs, it is important for researchers to probe more deeply into how different student-level variables can affect students' academic achievement in DLI programs involving other language pairs, in particular those that involve non-alphabetic languages. Understanding the effectiveness of such programs is crucial in that it will allow us to consider the effectiveness of DLI program in a broader educational landscape (Lee & Wang, 2021). Since the pathways to bilingualism and biliteracy in these languages may be distinct, crossing these linguistic boundaries may lead to a more nuanced understanding of learning and instruction.

Given this backdrop, the current study adds to the small yet much-needed literature of studies examining the effect of Mandarin DLI education on student achievement. Using a longitudinal dataset, this paper investigated how math and ELA scores grew over 3 academic years for linguistically diverse students in Mandarin DLI. We used growth curve modeling (GCM) to analyze student trends in math and ELA scores over time and variations in changes over time among individual students in elementary school. By focusing on data from one school over three years, the current study investigated the following questions:

1. As a group, did students' academic performance, indicated by their scores on state-mandated tests on math and ELA significantly improve over the three years?
2. Did students' academic performance vary over time as a function of their race and language backgrounds?

3 Methodology

3.1 *The Research Setting*

This study reports longitudinal data on 259 students enrolled in a public school that offered whole-school Mandarin immersion.¹ The school is situated in an urban setting in California and enrolls families of diverse backgrounds. In order to ensure its diversity, the school uses a categorical lottery preference of 20% for economically

¹ In the U.S., whole-school immersion describes a situation when the entire school offers an immersion program in one language. Some schools offer immersion programs in multiple languages. There are also schools which offer both a traditional English-only program as well as an immersion program.

Table 1 Participants by grade level

Year	3rd Grade (n)	4th Grade (n)	5th Grade (n)	6th Grade (n)	7th Grade (n)
2016–2017	54	45	50	35	
2017–2018	52	54	42	29	28
2018–2019		49	52	28	18

Table 2 English language acquisition status

Year	EL (n, %)	IFEP (n, %)	RFEP (n, %)	EO (n, %)	Missing (n, %)
2016–2017	19 (7.3)	42 (16.2)	7 (2.7)	116 (44.8)	75 (29.0)
2017–2018	5 (1.9)	41 (15.8)	25 (9.7)	134 (51.7)	54 (20.8)
2018–2019	1 (.4)	21 (8.1)	23 (8.9)	102 (39.4)	112 (43.2)

Note: *EL* English learners, *IFEP* Initial Fluent English Proficient, *RFEP* Reclassified Fluent English Proficient, *EO* English Only

disadvantaged students. Currently, according to its website, the school enrolls: 48% students of Asian background, 30% students of two or more backgrounds, 8% White, 5% African American, and 4% Hispanic Latino. In Kindergarten to second grade, 90% of instruction is conducted in Mandarin for Chinese language arts (CLA), math, social studies, art, and science and 10% in English for ELA, physical education (P.E.), and music in K-2; in Grades 3–4, 70% of the instruction is conducted in Mandarin for CLA, math, and science, while 30% is in English for ELA, science, history/social studies, art, music, and P.E. The proportions of instructional time for the same subjects reach 50% for both instructional languages in Grades 5–6.

Participant demographics are presented in Table 1. Data were collected across three academic years: 2016–2017 ($N = 184$), 2017–2018 ($N = 205$), and 2018–2019 ($N = 147$). A total of 259 unique cases were included across the period of data collection. Table 2 reports participants’ English Language Acquisition status by academic year.

Of the 259 participants, 156 were reported to speak English as the primary home language; 77 were reported to speak Chinese or Mandarin Chinese; 1 was reported to speak Spanish; and home language data were missing for 25 students. Race for participants was as follows: White: 32, Black: 8, Asian: 108, 2 or More Races: 0, Hispanic: 19, Missing: 92 (Total $N = 259$). Data on students’ English language acquisition status is reported in Table 2.

3.2 Data Source and Analyses

The present study analyzed longitudinal, de-identified data from students’ scores on the Smarter Balanced Summative Assessments for math and ELA beginning in Grade 3 and provided for three consecutive academic years. Further information on the Smart Balance Summative Assessments (SBAC) can be found on the website of California Department of Education (<https://www.cde.ca.gov/ta/tg/sa/sbacsummative>).

asp). Assessment data were collected during the spring semester of each academic year. In this study, data collected from individuals at three time points were used to analyze trends over time and variations in changes over time among individuals. Specifically, we were interested in examining the longitudinal relations between students' demographic information (i.e., race and language background) and their academic achievement across three waves of measurement. The categorical predictor, race, was coded as 0 = White, 1 = Black/African American, 2 = Asian, 3 = Hispanic/Latino, and 4 = Two or more races. The language backgrounds were coded as English = 0, Unidentified = 1, Spanish = 2, Chinese (Mandarin) = 3.

Growth curve modeling (GCM) was used to explore students' growth trajectory of academic performance over time. As a widely used method to analyze longitudinal data from repeated measures, GCM captures the collection of individual trajectories over time and demonstrates inter-individual variability in intra-individual patterns of change over time (Curran et al., 2010).

In the analysis, we used a growth curve model with two levels and two predictors. Each participant was observed on the variable Y at three different times. Conceptualized as a multilevel model, the variable time is a level 1 variable. Time was coded 0, 1, and 2. The intercept is the predicted value when time is 0. Each subject has its own intercept and slope, expressed as random effects at level 2. We can write this model using multiple equations, as shown here (Frey, 2018).

$$\text{Level1 : } Y_{ij} = \beta_{0j} + \beta_{1j} \text{Time} + r_{ij}.$$

$$\text{Level2 : } \beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{00} + u_{1j}$$

At level 1, Y_{ij} represents the outcome Y for level one unit i nested in level 2 unit j and is equal to a level one intercept β for predicted value when time is 0, and r_{ij} presents residual or unexplained variance for level 1 unit i nested in level 2 unit j . At level 2, the level 1 intercept, β_{0j} , is set as the outcome in a new regression equation with two components: the level 2 intercept, γ_{00} , and a random parameter, u_{0j} , which is the level 2 residual variance. The level 2 random parameter, u_{0j} , is what allows the model to vary by the higher-level unit. Student-level variables, including race and home language were included in the analysis as predictors. The path models used for GCM in our study are displayed in Figs. 1 and 2.

4 Results

Descriptive statistics, correlations and analysis of variance (ANOVA) were conducted in SPSS. *Mplus* 8.2 (Muthén & Muthén, 1998–2017) was selected to perform growth curve analysis due to its powerful function in modeling latent variables

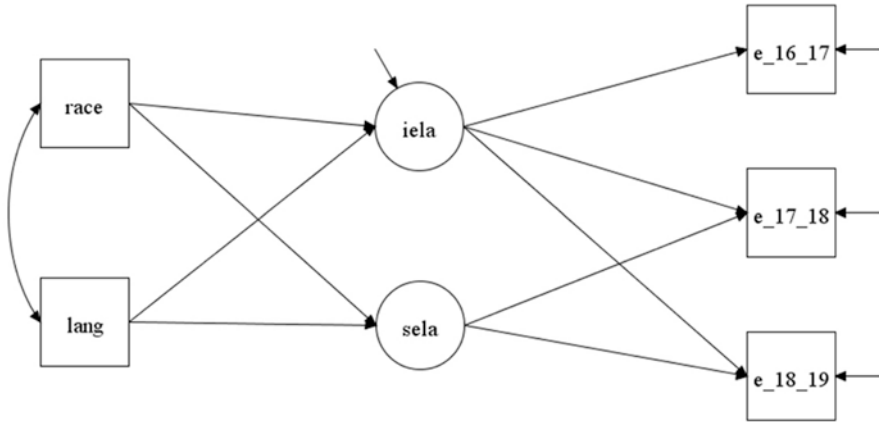


Fig. 1 Path model for ELA

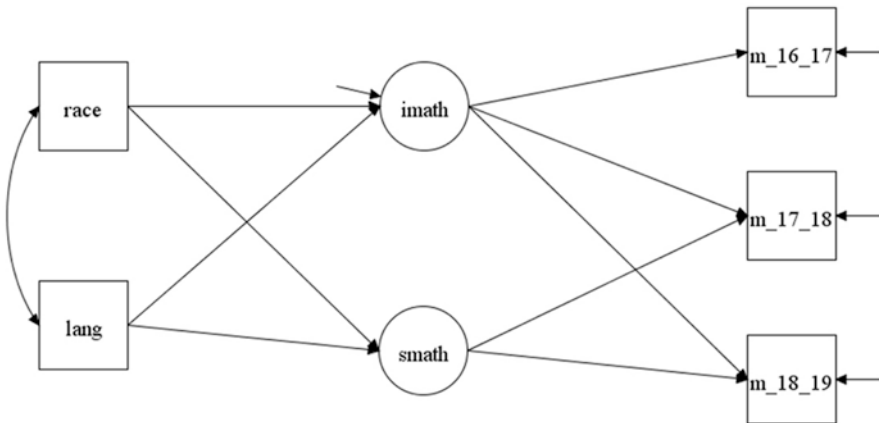


Fig. 2 Path model for math

and its capacity to estimate categorical variables by using robust weighted least squares, such as WLSM and WLSMV (Finney & DiStefano, 2013).

Table 3 displays the students’ demographic information, and Table 4 reports students’

SBAC data. In addition, the correlations among variables are displayed in Table 5.

As a group, children showed growth over time, as reflected by increased mean scores on math and ELA assessments. Participants’ assessment scores were all strongly correlated within and across subjects at each data collection point (all correlations > .6, and all *ps* < .001). Students’ ELA classification was largely stable over time for a majority of participants, however, all students who were initially classified as ELs were reclassified between 2016 and 2019. Specifically, out of the

Table 3 SBAC students' demographic information

Characteristic	N	Percentage
Race	177	
White	41	23.16
Black/African American	9	5.08
Asian	115	64.59
Hispanic/ Latino	12	6.78
Language	257	
English	166	64.59
Chinese (Mandarin)	84	32.68
Spanish	1	0.39
Unidentified	6	2.33
Migrant (No)	259	100

Table 4 SBAC student assessment scores

Assessment	M (SD)	N
English Language Arts		
ELA 2016–17	2557.69 (101.83)	183
ELA 2017–18	2582.37 (89.70)	205
ELA 2018–19	2610.70 (81.02)	147
Mathematics		
Math 2016–17	2570.75 (75.32)	182
Math 2017–18	2597.90 (76.81)	205
Math 2018–19	2621.01 (73.30)	147

Table 5 Correlations among all variables

	ELA 16–17 (N)	ELA 17–18 (N)	ELA 18–19 (N)	Math 16–17 (N)	Math 17–18 (N)	Math 18–19 (N)
ELA16–17	–					
ELA17–18	.857 (146)	–				
ELA18–19	.790 (94)	.818 (147)	–			
Math16–17	.781 (181)	.731 (147)	.720 (94)	–		
Math17–18	.714 (146)	.767 (205)	.695 (147)	.777 (147)	–	
Math18–19	.660 (94)	.707 (147)	.783 (145)	.795 (95)	.772 (147)	–

19 total participants who were initially classified as ELs, 10 (52%) were reclassified to RFEP, 3 (15%) were reclassified to EO, 1 (5%) was reclassified to IFEP, and subsequent data were missing for 5 (26%).

To answer the first question, a repeated measures ANOVA was used to determine if any significant differences existed in students' math and ELA scores across the three time points. Cases with missing values for any one of the three time points were excluded from the analysis, resulting in $N = 95$ for math and $N = 94$ for ELA, respectively.

For math scores, Mauchly’s Test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 7.729, p = .021$, therefore degrees of freedom were corrected using the Huynh-Feldt correction ($\epsilon = .943$). Our results show that students’ math scores differed significantly between the three time points $F(1.887, 177.289) = 176.729, p < 0.001$. Post-hoc analysis with a Bonferroni adjustment revealed that students’ math scores were significantly higher at Time 2 (2017–2018) than Time 1 (2016–2017) (Mean difference = 54.926, 95% CI [41.552, 68.299], $p < .001$); Time 3 (2018–2019) was also significantly higher than Time 2 (Mean difference = 43.117, 95% CI [29.278, 56.956], $p < .001$) and Time 1 (Mean difference = 98.043, 95% CI [87.230, 108.855], $p < .001$).

For ELA scores, Mauchly’s Test indicated that the assumption of sphericity had been violated, $\chi^2(2) = 194.027, p < .001$; therefore, degrees of freedom were corrected using the Greenhouse-Geisser correction ($\epsilon = .532$). Our results show that, as a group, students’ ELA scores differed significantly between the three time points $F(1.065, 98.952) = 17.015, p < .001$. Post-hoc analysis with a Bonferroni adjustment revealed that their ELA scores were significantly higher at Time 2 (2017–2018) than Time 1 (2016–2017) (Mean difference = 79.223, 95% CI [12.211, 146.235], $p < .015$); Time 3 (2018–2019) was also significantly higher than Time 2 (Mean difference = 54.638, 95% CI [40.367, 68.910], $p < .001$) and Time 1 (Mean difference = 133.862, 95% CI [64.573, 203.150], $p < .001$).

To summarize, our results show that, as a group, students in this focal Mandarin immersion program made significant, positive changes in their math and ELA scores on state-mandated tests across the three academic years under investigation.

Research question 2 asked if students’ academic performance varied over time as a function of their race and language backgrounds. The latent growth curve (LGC) model was used to explore the longitudinal growth of math and ELA on demographic variables. Figures 3 and 4 display the longitudinal plots for math and ELA scores, respectively.

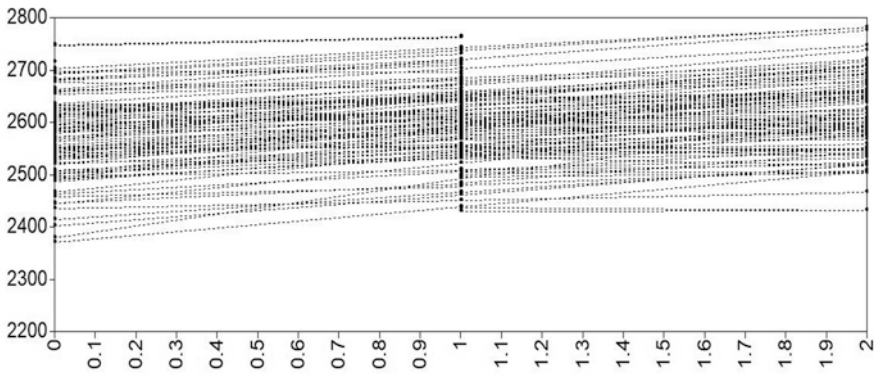


Fig. 3 Trajectories of math achievement scores across 2016–17, 2017–18, and 2018–19. The x-axis indicates the three time points in our analysis. The y-axis is students’ math scores

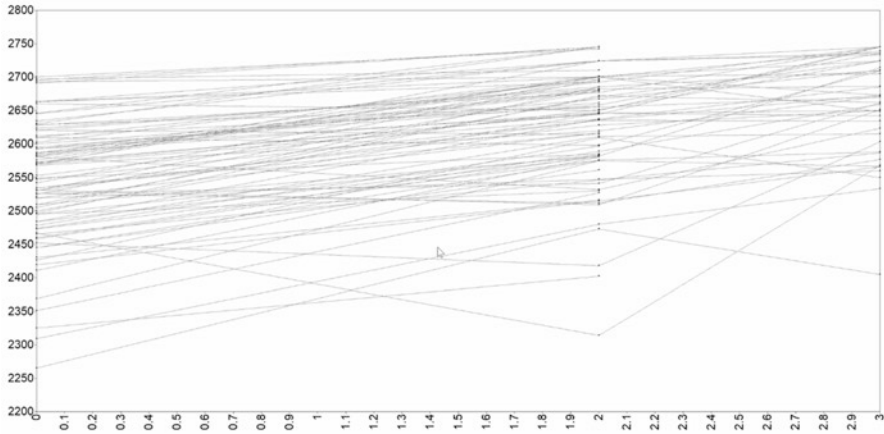


Fig. 4 Trajectories of ELA achievement scores across 2016–17, 2017–18, and 2018–19. The x-axis indicates the three time points in our analysis. The y-axis is students’ ELA scores

Table 6 Model fit information for ELA literacy and mathematics

	ELA	Mathematics
Chi-Square Test of Model Fit		
Value	11.89	7.91
Degrees of Freedom	5	5
P-Value	.03	.00
RMSEA		
Estimate	.04	.05
90 Percent C.I.	.02, .16	.00, .19
CFI	.972	.972
TLI	.949	.950
SRMR	.08	.04

The model fit information for ELA and math is shown in Table 6. For the GCM with math, the CFI and TLI are .972 and .950, respectively, indicating a good model fit. For the GCM with ELA, the CFI and TLI are .972 and .949, respectively, indicating a good model fit as well.

Table 7 shows the latent growth curve model results for math and ELA. When race and language are held constant at 0, the intercepts for math and ELA are 2500.70 and 2568.42, respectively. The variances of the intercepts for the math and ELA assessments are 53.74 and 43.10; however, both the variances of the slope for time of math and ELA scores are 0. The LGC model results also show residual variance for each time point (math and ELA scores from 2016–2017, 2017–18, and 2018–19), which suggests that the students’ math and ELA scores varied significantly across the three years. However, though students’ math and ELA scores

Table 7 Model results for ELA literacy and mathematics

	Mathematics		English Language Arts
	Estimate		Estimate
IMATH on		IELA on	
Race	.543	Race	12.752
Lang	.102	Lang	-7.049
SMATH on		SELA on	
Race	-.130	Race	3.20
Lang	0.306	Lang	1.41
IMATH with		IELA with	
SMATH	0.000	SELA	0.000
Intercepts		Intercepts	
M_16_17	.000	E_16_17	.000
M_17_18	.000	E_17_18	.000
M_18_19	.000	E_18_19	.000
IMATH	2500.70	IELA	2568.42
SMATH	53.74	SELA	43.10
Residual variances		Residual variances	
M_16_17	975.82	E_16_17	1497.085
M_17_18	2917.04	E_17_18	785.812
M_18_19	1652.35	E_18_19	1375.63
IMATH	4714.57	IELA	8737.45
SMATH	.000	SELA	.000

varied significantly at each time of measurement, there is no random slope variance. This finding suggests that race and language background had no influence on students’ trajectory growth change in math or ELA scores, because every student’s trajectory was the same.

In summary, our analyses found that, as a group and across three academic years, students in this focal Mandarin DLI program made significant academic progress indexed by their scores on math and ELA, and no difference in their performance was found across the three years as a function of their race and language backgrounds, thus indicating that students in this program, regardless of their race and language backgrounds, progressed equally well.

5 Discussion

The primary purpose of this research was to investigate DLI students’ longitudinal growth trajectories and evaluate the role that student characteristics played in learning. The students represented a diverse sample of students attending a public DLI program who varied in race, home language background, and English language proficiency.

Consistent with previous samples from other language pairs (e.g., Steele et al., 2017; Watzinger-Tharp et al., 2018), the participants in this Mandarin DLI program demonstrated significant growth in academic performance over time. Critically, longitudinal models indicated that students' math and ELA assessment scores changed in the same way over time with no effect of student-level variables on the slope of GCM models. This finding suggests that students' race and language background did not have any significant influence on the trajectory of growth change in math and ELA over time, or, in other words, racially and linguistically different subgroups of students in this study benefited equally from the Mandarin immersion program.

The fact that all students demonstrated positive growth trajectories regardless of race or language background has important implications for the field because it suggests that Mandarin DLI education, despite teaching school subjects in two highly contrasting languages, still provides an effective pedagogical model for children from diverse backgrounds to learn school subjects. Our analysis contributes to the growing literature by adding evidence coming from a Mandarin immersion program.

These data stand in contrast to national norms that have shown persistent gaps in math and reading scores based on students' language background. In particular, students who are identified as ELs tend to have scores on National Assessment of Educational Progress (NAEP) math and reading assessments that are consistently below students who are not classified as ELs (Nation's Report Card, 2021). Thus, when considering the full education landscape, the current study is important because it suggests that DLI programs – including Mandarin DLI – offer a potential solution for the persistent gaps observed nationally. Additionally, contrary to the criticism of being “elitist” (e.g., Peel, 2019) that Mandarin DLI programs often receive, the current study demonstrates that it is a viable and equitable program model serving language-minority as well as language-majority students.

Reflecting on the main theme of the volume, as well as our own experiences carrying out this project, albeit small scale, the importance of boundary crossing, pioneered and exemplified by Tucker, is becoming ever clearer. Researchers and practitioners of Chinese as a second language are often siloed into fields bounded by the language as well as organizational affiliation; however, learners of the Chinese language are becoming younger (Li, 2018), and Chinese language programs are offered increasingly more through mainstream schools (as opposed to weekend community schools; e.g., Lü, 2014, 2019) at the precollegiate, especially elementary school levels (American Councils for International Education, 2017). Because this trend is expected to continue, it is all the more important for researchers and practitioners to cross the boundaries of language, research disciplines, as well as research methodologies to conduct investigations that not only build on the legacy of trailblazers like Tucker, but also provide insight into pedagogies that promote learning as DLI programs grow to serve more students from culturally, linguistically, and socioeconomically diverse backgrounds.

6 Limitations and Conclusion

Our study demonstrates that Mandarin immersion is viable in that all students in our sample, regardless of language background or race, showed significant academic progress over the years in the focal Mandarin immersion program, which is a public school serving a diverse population. Our study confirms the currently held belief that DLI education, if appropriately implemented, is “the most powerful school reform model for high academic achievement” (Thomas & Collier, 2012, p. 6), with additional data from a Mandarin immersion program. Our focal school offers an excellent model of centering diversity, equity, inclusivity as well as academic excellence through Mandarin immersion. However, the current dataset is relatively small, only includes data over three years, and reflects disproportionate enrollment of Asian and non-Hispanic White students relative to Black and Hispanic students. Future research should work with schools and school districts with larger numbers of Mandarin DLI programs, ideally at multiple sites, and extend longitudinal investigations on larger data sets over longer periods of time.

Despite the limitations, our study suggests that Mandarin DLI is a viable program model with the potential to deliver the benefits of bilingualism and biculturalism. However, positive impacts of DLI education need to be extended to all students not only by implementing more DLI programs across the nation, but also by ensuring that students “in all ZIP codes have the opportunity to formally study languages other than English through the more equitable distribution of dual-language programs” (Flores, 2019, para. 2). We are optimistic that efforts to expand the high-quality implementation of immersion education programs will be pivotal for meeting the needs of an increasingly diverse student population.

References

- Alvear, S. A. (2019). The additive advantage and bilingual programs in a large urban school district. *American Educational Research Journal*, 56(2), 477–513.
- American Councils for International Education. (2017). *The national dual language immersion research alliance: Where practitioners and researchers align*. Retrieved on June 2, 2018, from <https://www.americancouncils.org/sites/default/files/DLI-portfolio.pdf>
- Baig, F. (2011). *Investigating the motivations of parents choosing language immersion education for their child*. Unpublished doctoral dissertation. University of Iowa.
- Bibler, A. (2021). Dual language education and student achievement. *Education Finance and Policy*, 16(4), 634–658.
- Bruck, M., Jakimik, J., & Tucker, G. R. (1971). Are French immersion programs suitable for working-class children? A follow-up investigation. *Word*, 27(1–3), 311–341.
- Campbell, R. N. (1984). The immersion education approach to foreign language teaching. In *Studies on immersion education: A collection for United States educators* (pp. 114–143). California State Department of Education.
- Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognitive Development*, 11, 121–136.

- de Jong, E. J. (2002). Effective bilingual education: From theory to academic achievement in a two-way bilingual program. *Bilingual Research Journal*, 26(1), 65–84.
- Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 439–492). Information Age Publishing.
- Flores, N. (2019). How to dismantle elite bilingualism. *Education Week*. Retrieved on June 20th, 2021, from: <https://www.edweek.org/teaching-learning/the-truth-about-bilingualism-its-only-for-some-students/2019/01>
- Fortune, T. W., & Ju, Z. (2017). Assessing and exploring the oral proficiency of young Mandarin immersion learners. *Annual Review of Applied Linguistics*, 37, 264–287.
- Fortune, T. W., & Song, W. (2016). Academic achievement and language proficiency in early total Mandarin immersion education. *Journal of Immersion and Content-Based Language Education*, 4(2), 168–197.
- Fortune, T. W., & Tedick, D. J. (2003). *What parents want to know about foreign language immersion programs*. ERIC Digest. ERIC Clearinghouse on Languages and Linguistics. Retrieved August 5, 2016, from <https://www.ericdigests.org/2004-4/parents.htm>
- Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vols. 1–4). SAGE Publications.
- Genesee, F., & Fortune, T. (2014). Bilingual education and at-risk students. *Journal of Immersion and Content-Based Language Education*, 2, 196–209.
- Genesee, F., Tucker, G. R., & Lambert, W. E. (1975). Communication skills of bilingual children. *Child Development*, 46(4), 1010–1014.
- Howard, E. R., Lindholm, L. K., Rogers, D., Olague, N., Media, J., Kennedy, B., ... & Christian, D. (2018). *Guiding principles for dual language education*. The Center for Applied Linguistics (CAL).
- Jaumont, F. (2017). *The bilingual revolution: The future of education is in two languages*. TBR Books.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Newbury House Publishers.
- Lambert, W. E., Tucker, G. R., & D'Anglejan, A. (1973). Cognitive and attitudinal consequences of bilingual schooling. *Journal of Educational Psychology*, 65(2), 141–159.
- Lee, J. S., & Wang, T. (2021). A review of Korean/English and Mandarin/English dual language programs in the United States. *Language Teaching for Young Learners*, 3(1), 28–65.
- Lenker, A., & Rhodes, N. (2007). Foreign language immersion programs: Features and trends over thirty-five years. *ACIE Newsletter/The Bridge: From Research to Practice*, 1–8.
- Li, Y. M. (2018). Some thoughts on the younger-age trend of overseas Chinese learners. *Chinese Teaching in the World*, 3, 291–301.
- Lindholm-Leary, K., & Hernández, A. (2011). Achievement and language proficiency of Latino students in dual language programmes: Native English speakers, fluent English/previous ELLs, and current ELLs. *Journal of Multilingual and Multicultural Development*, 32(6), 531–545.
- Lü, C. (2014). Chinese heritage language schools in the United States. In X. L. Curdt-Christiansen & A. Hancock (Eds.), *Learning Chinese in diasporic communities many pathways to being Chinese* (pp. 81–96). John Benjamins Publishing.
- Lü, C. (2019). *Chinese literacy learning in an immersion program*. Palgrave Macmillan.
- Lü, C. (2020). Teaching and learning Chinese through immersion: A case study from the North American context. *Frontiers of Educations in China*, 15, 99–141.
- Marian, V., Shook, A., & Schroeder, S. R. (2013). Bilingual two-way immersion programs benefit academic achievement. *Bilingual Research Journal*, 36(2), 167–186.
- Morita-Mullaney, T., Renn, J., & Chiu, M. M. (2021). Contesting math as the universal language: A longitudinal study of dual language bilingual education language allocation. *International Multilingual Research Journal*, 15(1), 43–60.
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide*. Muthén & Muthén.

- Nation's Report Card. (2021). *National student group scores and score gaps*. Retrieved on July 10th, 2021 from: <https://www.nationsreportcard.gov/reading/nation/groups/?grade=8>
- Palmer, D. K. (2010). Race, power, and equity in a multiethnic urban elementary school with a dual immersion "strand" program. *Anthropology & Education Quarterly*, 41(1), 94–114.
- Palmer, D. K., Cervantes-Soon, C., Dorner, L., & Heiman, D. (2019). Bilingualism, biliteracy, biculturalism, and critical consciousness for all: Proposing a fourth fundamental goal for two-way dual language education. *Theory Into Practice*, 58(2), 121–133.
- Peel, S. (2019). Portland public schools parents wage war as international travel for students may get its wings clipped. *Willamette Week*. Retrieved on July 10th, 2021 from <https://www.wweek.com/news/schools/2019/02/20/portland-public-schools-parents-wage-war-as-international-travel-for-students-may-get-its-wings-clipped/>
- Rubio, F. (2018). Language education in elementary schools: Meeting the needs of the nation. *Foreign Language Annals*, 51(1), 90–103.
- Serafini, E. J., Rozell, N., & Winsler, A. (2022). Academic and English language outcomes for DLLs as a function of school bilingual education model: The role of two-way immersion and home language support. *International Journal of Bilingual Education and Bilingualism*, 25(2), 552–570.
- Steele, J. L., Slater, R. O., Zamarro, G., Miller, T., Li, J., Burkhauser, S., & Bacon, M. (2017). Effects of dual-language immersion programs on student achievement. *American Educational Research Journal*, 54(1_suppl), 282S–306S.
- Thomas W. P., & Collier, V. P. (2002). *A national study of school effectiveness for language minority students' long-term academic achievement*. UC Berkeley: Center for Research on Education, Diversity and Excellence. Retrieved on June 20th, 2021 from <https://escholarship.org/uc/item/65j213pt>
- Thomas, W. P., & Collier, V. P. (2012). *Dual language education for a transformed world*. Fuente Press.
- Tucker, G. R. (1979). Bilingual education: Some perplexing observations. *Educational Evaluation and Policy Analysis*, 1(5), 74–75.
- U.S. Census Bureau. (2015). Detailed languages spoken at home and ability to speak English for the population 5 years and over for United States: 2009–2013, American Community Survey. Retrieved on May 20th, 2021 from <https://www.census.gov/data/tables/2013/demo/2009-2013-lang-tables.html>
- Valdés, G. (1997). Dual-language immersion programs: A cautionary note concerning the education of language-minority students. *Harvard Educational Review*, 67(3), 391–429.
- Watzinger-Tharp, J., Swenson, K., & Mayne, Z. (2018). Academic achievement of students in dual language immersion. *International Journal of Bilingual Education and Bilingualism*, 21(8), 913–928.
- Watzinger-Tharp, J., Tharp, D. S., & Rubio, F. (2021). Sustaining dual language immersion: Partner language outcomes in a statewide program. *The Modern Language Journal*, 105(1), 194–217.

Evaluating the Modular Curriculum of Chinese Language in Singapore Primary Schools: Insights from Students and Teachers



Dongbo Zhang , Shouhui Zhao, and Xiaoxi Sun

Abstract Since Singapore became an independent republic, careful language-in-education planning that caters to the economic, social, and political development of the country has never been abated. A notable case in point is the constant reviews and reforms of the curriculum of Chinese language (CL), a school subject required of ethnic Chinese children, to respond to gradual home language shift from CL toward English, which is the medium of instruction, in the country. In 2008, a differentiated Modular Curriculum (MC) began to be implemented in all primary schools. We were subsequently commissioned by the Ministry of Education to evaluate the MC. We analyzed CL teaching materials, observed and coded CL classes, and engaged students and CL teachers through various methods. In this chapter, based on student surveys as well as a teacher survey and focus group discussions, we report some evaluation findings on how students' interest in CL learning and use changed, how the MC and its underlying principles were perceived by teachers, and what difficulties and challenges teachers experienced under the MC. Based on the findings, we discuss the interface between sociolinguistics, curriculum innovation and reform, and language policy and planning; and underscore boundary crossing in curriculum and program evaluation toward evidence-based language-in-education planning.

D. Zhang (✉)

School of Education, University of Exeter, Exeter, UK
e-mail: d.zhang4@exeter.ac.uk

S. Zhao

Department of Foreign Languages, University of Bergen, Bergen, Norway
e-mail: Shouhui.Zhao@uib.no

X. Sun

Department of Languages, Cultures and Visual Studies, University of Exeter, Exeter, UK
e-mail: X.Sun3@exeter.ac.uk

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_11

231

Keywords Chinese language education · Curriculum and program evaluation · Bilingual education · Language-in-education planning · Boundary crossing · Methodological pluralism · Educational stakeholders

1 Introduction

In bi-/multi-lingual societies, particularly post-colonial societies, where there is often a contentious agenda on balancing the promotion of the colonial language and the maintenance of an ethnic language and cultural heritage, school curriculums are constantly reviewed and reformed to meet with the realities of the evolving sociolinguistic landscape. Singapore, a multilingual country with the Chinese as the largest ethnic group, is no exception. Since Singapore became an independent republic, careful language-in-education planning that caters to its economic, social, and political development has never been abated. A notable case in point is the constant reviews and reforms of the curriculum of Chinese language (CL), a school subject required of ethnic Chinese (EC) children, to respond to gradual home language shift from CL toward English, which is the medium of school instruction, in the country. After a two-year pilot, a Modular Curriculum (MC), which was developed by the Ministry of Education (MOE) in response to the recommendations of the Chinese Language Curriculum and Pedagogy Review Committee (CLCPRC), was launched in 2008 in all primary schools to cater to the different learning needs and CL abilities of children from different home language backgrounds.

To evaluate the MC, the MOE commissioned the CL research team, headed then by the first two authors in the Center for Research in Pedagogy and Practice (CRPP), National Institute of Education (NIE), to conduct a multi-year project about 3 years after the MC's official launch. As part of the evaluation project, we analyzed CL teaching materials, observed CL classes, and studied students and teachers through various methods. In this chapter, based on student questionnaires as well as teacher questionnaires and focus group (FG) discussions, we report some evaluation findings on how the interest in CL learning and use changed in students from different home language backgrounds, how the MC and its underlying principles were perceived by teachers, and what difficulties and challenges teachers experienced under the MC. Based on the findings, we discuss the interface between sociolinguistics, curriculum and program innovation and reform, and language policy and planning; and underscore boundary crossing in curriculum and program evaluation toward evidence-based language-in-education planning.

2 Changing Landscape of CL Education in Singapore

Singapore is a multilingual, multi-ethnic country in Southeast Asia. A former British colony, Singapore became an independent republic from Malaysia in 1965. Bilingual education is the cornerstone of the educational system in Singapore, where there are four official languages including English as well as the ethnic languages of the three major ethnic groups (i.e., Chinese, Malay, and Indian), that is, Chinese, Malay, and Tamil. English, in addition to being a school subject itself, is the medium of instruction (and also the *de facto lingua franca* in the country) and the three ethnic languages are designated mother tongue languages (MTLs) of the ethnic groups and learned by the respective group as a school subject. For example, CL is the MTL of ethnic Chinese, who are also the largest ethnic group in the country (about 75% of the population). Singaporean students are taught to become bilingual in English as well as the MTL. In the educational discourse in Singapore, English is often referred to as the “first language” of Singaporeans and the MTL as the “second language.” This designation, however, does not reflect the conventional sense of first (L1) versus second language (L2) in the second language acquisition (SLA) literature, that is, L1 being the native or home language and L2 an additional language acquired temporally after L1. Rather, it is based on the importance ascribed to English and the MTL in the society and school curriculum. English is thus sometimes described as the “first school language” whereas the MTL is the “second school language” in Singapore (Pakir, 1992).

The designation of CL as EC students’ *mother tongue* does not suggest that CL is necessarily their L1. Not all EC children grow up speaking CL. In fact, as a result of the global influence of English and the importance ascribed to English in the society and in schools, an increasing number of EC children use English as their only or dominant home language (about 40% of those entering primary school, as reported in CLCPRC, 2004; see also Zhao & Liu, 2010). This home language shift from CL to English has had strong ramifications on the sociolinguistic landscape and CL education in Singapore. In both media and scholarly publications, there are discussions on or concerns about how language profiles of school children are changing and how school curriculum should be reformed to accommodate those changes and revert declining motivation in students for learning and using CL.

CL curriculum and teaching in Singapore was long influenced by a so-called L1 approach, where a strong emphasis was placed on the development of a high level of literacy, including an ability to recognize as well as write a large number of Chinese characters. Chinese characters, which are based on strokes and fundamentally different from alphabetic writing systems such as English (see Zhang, 2017), are often cited by language educators and students to be a particularly challenging aspect of learning Chinese as an additional language (Hu, 2010). While the L1-based approach justifiably characterizes primary school education in China, where children largely grow up speaking CL and have a good command of CL oral proficiency upon entering primary school, this approach and a similar, mandated goal for literacy, particularly writing (which relies on stroke memorization), for all students

failed to capture the local reality of English-medium, bilingual education, particularly the diverse range of language profiles and CL proficiency in school children. A lack of oral proficiency in CL in children from families that use English as the only or dominant language, for example, poses constraints to their CL learning in the early years of primary school, and an equally strong emphasis on character recognition and writing in CL teaching without aiming to build up oral proficiency can be demotivating to those students (CLCPRC, 2004). In short, the L1-based approach turned out unaccommodating in Singapore.

The foregoing discussion was the backdrop of the formation of the CLCPRC in 2004, which was charged by the MOE to review the teaching and learning of CL in schools and make recommendations for reform and innovation.¹ In a report released later in the year by the CLCPRC, it was recognized that a large majority of parents felt it important for children to study CL; nevertheless, more and more children entering primary school spoke predominantly English at home; and there would be a continuing trend of Primary/Grade 1 (P1) children having had little exposure to CL. The report underscored that it was unrealistic to expect most students to be equally proficient in both English and CL, and CL education should, more pragmatically, aim to “stimulate an interest in the language in all CL students and motivate them to use it long after they leave school” (CLCPRC, 2004, p. ii).

Among the many recommendations, which covered flexibility in curriculum, alternative assessments, creative teaching methods (e.g., using modern information and communication technology or ICT), teaching materials, and support for teachers, two are particularly noteworthy. First, it would not be realistic to require all students to write characters (or script-writing as called in the CLCPRC report) and use them at the same time they learn to recognize them. Placing an equal emphasis on script-writing, which was usually taught through the traditional approach of copying and dictation (Liu et al., 2006), and character recognition, could be demotivating. The CLCPRC recommended that character teaching adopt the “*Recognize First, Write Later*” principle such that students are taught to recognize a large number of characters initially for promoting early meaningful reading and building reading interest; script-writing can then be delayed to allow more time for character recognition and reading activities, and character use, such as for composition purposes, can be aided with the use of ICT tools. Second, for the majority of students, emphasis should be on supporting the development of listening and speaking (and reading) for functional communication purposes. More competent students can be supported in developing all four skills. The CLCPRC consequently recommended that a customized curriculum comprised of Bridging, Core, and Enrichment modules be developed to provide a flexible approach such that all students, through learning with differentiated objectives, are supported to achieve their best learning potential and develop lasting interest in learning and using CL in school as well as after leaving school.

¹ This was not the first time that CL education was reviewed in the country. Prior to this review, two others had been conducted in 1999 and 1992, respectively (Chin, 2018).

Those recommendations were accepted by the MOE, and the Curriculum Planning and Development Division (CPDD) of the MOE subsequently started to develop a modular curriculum and aimed to pilot it with P1 and P2 students in 2006. According to the 2007 Chinese Language Syllabus (Primary) (MOE, 2006), students are assigned based on CL proficiency to study either *Gaoji Huawen* (Higher Chinese; for more competent students) or *Huawen* (Chinese; for less competent students, esp. those from English-dominant homes). Linking these two streams is a Core Module that every student must take (70–80% of instructional time). Depending on CL proficiency, some students also study, with 20–30% of instructional time, either a Bridging/Reinforcement Module (which is preparation for the Core Module and where a strong focus is placed on listening and speaking), whereas others also study an Enrichment Module (which is an extension of the Core Module and where there is a strong emphasis on writing and composition).

To support the development of the curriculum and its pilot, the MOE commissioned the CL research team at CRPP, NIE to survey home language use in kindergarten children and conduct a corpus-based study, based on classroom observation of children's CL use as well as elicitation tasks, to generate oral vocabulary lists based on children from different home language backgrounds (Zhao et al., 2007; see also Goh, 2017). The team was also subsequently commissioned to conduct a classroom observation study during the pilot of the MC in 2006 and 2007 with P1 and P2 students in 16 primary schools (Liu & Zhao, 2008). Students in pilot schools were found to be more engaged in learning; classes in those schools were characterized by greater percentages of student-centered activities (e.g., oral presentation and group work) and much less individual seatwork (e.g., character copying and worksheets). A much greater proportion of activities characterized by teacher-student or student-student interaction was found in classes studying the Bridging module than in those studying the other two modules. The MC was formally launched in 2008 in all primary schools. After about 3 years of its national, full-fledged implementation, the CL research team was commissioned by the MOE again to conduct a large-scale evaluation of it, which we describe in detail later.

3 Language Curriculum/Program Evaluation and Boundary Crossing

There has long been the question of *what works* in language education. To answer this question, curriculums and programs need to be rigorously evaluated. “To date, however, program evaluation largely has been ignored by the mainstream of applied linguistics, and as a result the capacity of evaluation to transform how we inquire, reason, and act in relation to language programs is yet to be realized” (Norris, 2016, p. 169). Evaluation aims to uncover the multiplicity of the truth about a program, and involves “the gathering of information about any of the variety of elements that constitute educational programs, for a variety of purposes that primarily include

understanding, demonstrating, improving, and judging program value” (Norris, 2006, p. 579). Curriculum and program evaluation is often developmental in nature and entails a long-term goal toward program sustainability and evidence-based language-in-education planning and innovation (Donato & Tucker, 2010; Nation & Macalister, 2020; Norris, 2016).

In language program evaluation, there is often an underlying interest in applied linguistics to combine it with language assessment (e.g., Davis, 2013; Lynch, 2003; Ross, 2009). This interest or emphasis seems very reasonable since the effectiveness of a curriculum or program for language proficiency development serves as essential evidence to gauge the extent to which program goals have been achieved (i.e., administrators and/or teachers wish to know whether “things have worked”). In other words, there is a built-in goal of assessing learner competence, and its change over time, in curriculum and program evaluation. Nevertheless, (program) evaluation and (language) assessment are distinct concepts (Norris, 2016). Consequently, the emphasis on learner assessment begs at least two questions that require understanding and approaching program evaluation in broader as well as more nuanced ways.

To begin with, what constitutes evidence of curriculum or program effectiveness through assessing students? While effectiveness is often established through testing language competence such as linguistic knowledge and the four skills, sustainable program development importantly necessitates contextualized understandings about students and their individual differences that accommodate both the goals of proficiency development and experiential components of learning (Donato & Tucker, 2010). In the SLA literature, language learners are recognized to differ in background, motivation and interest, learning strategies, classroom engagement, and willingness to communicate in the target language, among many other factors, which all have strong implications for their L2 development or learning outcomes (e.g., Dörnyei, 2005; Hiver et al., 2020). It would thus be very limiting to evaluate program effectiveness by only testing language competence following an outcome-oriented approach without due attention to students’ individual differences and learning engagement (e.g., classroom participation). In other words, student assessment should adopt a more inclusive and a pragmatist approach where not only are language skills measured and achievement gains established in quantified terms (a positivistic view) but evidence is collected to understand who the learners are, how they perceive their learning experience, and how various program-related experiences influence their learning process and outcomes (an interpretivist view that considers learning in personalized and contextualized ways) (Donato & Tucker, 2010).

Additionally, does a focus on students and student learning provide sufficient evidence for curriculum and program evaluation? Although students are arguably a fundamental consideration in language education and program evaluation, there are a multitude of other stakeholders, including but not limited to teachers, administrators, communities, and parents (Donato & Tucker, 2010; Kiely & Rea-Dickins, 2005; Tucker, 2000a). Accordingly, evidence of varied types needs to be collected from these stakeholders to generate insights into the context, the process (including insiders’ perceptions, experience, and practices), as well as outcomes of program

implementation. Teachers notably play a fundamental role in educational delivery and innovation. They are agents of change, a determinant factor in student learning and program success and sustainability, and policy makers (Donato & Tucker, 2010; Menken & García, 2010). It is their day-to-day work in the classroom that shapes student experience, motivates learning, and promotes language development. And teachers' classroom practices are influenced by a constellation of factors related to who they are; how they view language, learning, and teaching; and how they interpret curricular goals and negotiate micro and macro policies (Borg, 2006, Menken & García, 2010; see also Part II, this volume). It is thus no surprise that curriculum and program evaluation almost always seeks to engage teachers as a key stakeholder by looking into their instructional processes, perceptions and beliefs, and challenges and needs through a wide range of methods such as classroom observation, discourse analysis, and surveys and interviews (Davis & McKay, 2018; Donato & Tucker, 2010; Menken & García, 2010).

Rigorous curriculum and program evaluation thus requires methodological pluralism and information from diverse sources and of diverse types (Davis & McKay, 2018; Norris, 2016). It is inherent in program evaluation, which usually entails multiple purposes and goals, that boundaries be crossed between research paradigms (i.e., pragmatism), methodological approaches (quantitative and qualitative), and methods; and between stakeholders or policy actors (see also Part IV, this volume). This view on language program evaluation in light of boundary crossing is also aligned with shifting and pluralistic views on the nature of language, language learning and use, and teaching in the literature on SLA, language teaching, and education policy and planning (see Zhang and Miller, this volume).

G. Richard Tucker, whom this volume aims to honor, is arguably a pioneer in program innovation and evaluation and has exemplified boundary crossing through his numerous projects and publications that laid the foundation for researching, understanding, and improving language education and policy through stakeholder engagements and methodological pluralism. The St. Lambert Experiment (Lambert & Tucker, 1972) and the Pittsburgh FLES program (Donato & Tucker, 2010) notably provided compelling evidence on how engaging stakeholders and listening to their voices are crucial for understanding language programs and evaluating their effectiveness and impact in micro (school) as well as macro (sociocultural and sociopolitical) contexts. In many ways, our MC evaluation project was influenced by Dick's insights and followed the path he and his collaborators set for applied linguists to cross boundaries in language program evaluation toward evidence-based policy and planning.

4 The Modular Curriculum Evaluation Project

The MC evaluation project was designed to address four overarching goals: (1) to examine how cohesively the pedagogical principles recommended by the CLCPRC are represented across documents of the MC, esp. textbooks; (2) to understand how

the principles are translated into classroom teaching; (3) to understand how teachers view the MC and its implementation; and (4) to assess the impact of the MC on students' self-perceived competence and interest in learning and using CL. These goals were established not only based on key considerations, discussed earlier, of curriculum and program evaluation but also through discussions with the MOE, the primary user of the evaluation findings. The findings aimed to help the MOE monitor MC implementation and support teachers and schools to deliver the curriculum.

At different stages of the project from 2010 to 2012, we compared MC textbooks against previous textbooks to investigate, for example, how Chinese characters were represented for recognition and writing with reference to the “*Recognize First, Write Later*” principle recommended by the CLCPRC. We observed over 50 P2 classes studying different modules in 20 primary schools, which were a stratified random sample considering linguistic profiles of students in different types of primary schools, and coded classroom instructional strategies and focuses. Questionnaires were also administered to students in the participating schools. We also conducted FGs with CL teachers in the participating schools and administered an online questionnaire that targeted all primary school CL teachers in the country.²

It is of course impossible to report all project findings in this chapter. We decided to limit the scope by focusing on some findings related to the third and the fourth goal. Specifically, this chapter aims to answer the following three questions.

1. Did student interest in learning and using CL increase over time? Did the interest, and its change, if any, over time, differ between students studying different modules?
2. How did teachers view the MC in light of its emphasis on a differentiated approach and the major pedagogical principles recommended by the CLCPRC?
3. What difficulties and challenges, if any, were experienced by teachers under the MC?

²The project did not involve any direct testing of students' CL competence. This was purposefully planned for two reasons, in agreement with the MOE. First, during the project period, the MC was implemented nation-wide in all primary schools (i.e., there would not be any proper control group). Thus, it would be impossible to compare CL competence or skill attainment in students under the MC against that of students under the old curriculum. Second, the goal of the MC (and the recommendations of the CLCPRC) was not to boost national achievements, benchmarked on those of any earlier cohorts of students, but to make CL learning interesting to all through setting more realistic and differentiated goals for students from different home language backgrounds.

5 Participants and Data Collection

5.1 Student Surveys

A questionnaire was first administered in October 2010 (Time 1) in 17 primary schools where there were 3188 CL-studying P1 students. The students were then in their second semester of P1 and had been studying CL in their respective school for about 9 months. The same questionnaire was administered again to the same students about a year later in August 2011 (second semester of P2; Time 2). They were presented bilingually in simple English and Chinese and administered in CL classes where teachers were asked to read questions aloud and help children complete each section with necessary explanations.

The questionnaire began with some items on the general background of students, including, for example, date of birth, gender, module attending, and home language use. This was followed by instructions for answering the rest of the questionnaire, which included 69 Likert-scale items (and three warm-up items). Each item included a brief statement related to CL. Students were asked to circle an answer, from Yes (5), Maybe (4), Sometimes (3), Maybe Not (2), and No (1), to indicate the extent to which they thought the item represented their situation. The items covered students' attitude toward and self-perceived competence in listening and speaking, reading, and writing. Additional items were included on frequency of different types of language use and learning of life values related to Chinese culture from CL textbooks. The questionnaire also included a section on students' willingness to communicate (WTC) in CL, which considered age-appropriate topics (e.g., self-introduction, storytelling, giving instructions on playing games) and different contexts of CL use (e.g., in versus outside class with familiar versus unfamiliar interlocutors).

This chapter focuses only on self-perceived competence in CL listening and speaking (four items; e.g., *I can understand if people talk to me in Chinese*), attitude toward CL listening and speaking (four items; e.g., *I find it interesting to talk with people in Chinese*), and WTC (24 items; e.g., *Outside class, I am willing to explain how to play a game to a friend in Chinese*). This is because a key consideration in the CLCPRC report and the MC was that all students be supported, through a differentiated approach, to develop lasting interest in learning and using CL, and because oral language is a strong instructional focus in early primary grades, especially for those studying the Bridging module.

For Time 1, a total of 2708 valid questionnaires were collected where 174 were completed by Bridging students and 1362 and 1172 respectively by Core and Enrichment students. For Time 2, 1087 valid questionnaires were returned among which 69, 409, and 609 were completed by Bridging, Core, and Enrichment students, respectively. A total of 399 students (35, 113, and 251 for the three modules, respectively) completed the questionnaire for both times and formed the dataset for the statistical analysis reported later in this chapter. Cronbach's α ranged from .705 to .956 for the three variables (i.e., ability, attitude, and WTC).

5.2 *Teacher Questionnaire*

An online questionnaire was administered in Chinese in February 2011 targeting all primary school CL teachers. A total of 311 valid questionnaires were received from teachers in 108 primary schools. The teachers (85% females) covered all six grades (P1–P6) and consisted of a range of age groups (11.6% 20–25; 26% 26–30; 23.5% 31–35; 15.4% 36–40; 14.8% 41–50; and 8.7% over 50 years). Most of them (56.9%) were in the early years of CL teaching (0–5 years); 19.6%, 9.3%, 6.4%, 0.6%, and 7.1% had taught CL for 6–10, 11–15, 16–20, 21–25, and more than 25 years, respectively. In terms of their experience teaching under the MC, 13.5% reported less than 1 year; 12.2% 1 year; 16.4% 2 years; 12.5% 3 years; 16.1% 4 years; and 29.3% 5 years or more.

The questionnaire was comprised of several sections that covered a number of issues such as general perceptions of CL teaching, learning, and assessment; ICT (e.g., perceived ease and usefulness of ICT and frequency of using ICT for different purposes) (see Zhang et al., 2014); Chinese character teaching and learning; and efficacy of using differentiation strategies to motivate and support student learning. Because of space limitations, we only focus on items that targeted general perceptions of the MC for accommodating diversities in students and those toward Chinese characters.

Teachers' general perceptions of the MC were measured by 15 items that touched on three issues (see Table 2), including the importance of oral language in CL teaching and learning (e.g., *Teachers should give students enough time for oral language practice*); the capability of the MC for accommodating students from different home language backgrounds (e.g., *The Modular Curriculum offers a differentiated approach for teachers to cater to different learning needs in students*); and efficacy of using strategies of differentiation to motivate and support student learning (e.g., *I know how to adjust teaching based on different abilities in students*). Perceptions toward Chinese characters covered three issues and consisted of 11 items (see Table 3), including cultural and life values (e.g., *Chinese characters should be considered as an integral aspect of Chinese culture*); importance of character writing in CL learning (e.g., *Writing characters facilitates the recognition of characters*); and perceived student interest in writing characters (e.g., *Students are enthusiastic when I ask them to practice writing characters*). Cronbach's α ranged from .660 to .808 for different sections.

5.3 *Teacher Focus Groups*

We conducted 13 FG interviews in May 2010 with 107 teachers from the 20 participating schools. Each FG consisted of around eight members and lasted for about an hour. Each session was moderated by an experienced research team member familiar with CL education in Singapore and facilitated by an assistant. They were

conducted in CL and audio recorded and transcribed verbatim for subsequent coding and thematic analysis.

The FGs aimed to collect detailed qualitative data via group discussions where teachers could demonstrate their understandings and share views and experiences related to the MC. Like the survey, the discussions were structured to cover a range of issues encapsulated in six topics, each discussed through a set of open-ended questions. In this chapter we focus on questions and discussions targeting the MC's differentiated approach, pedagogical principles recommended by the CLCPRC, and difficulties and challenges teachers experienced under the MC.

6 Findings

6.1 Students' Self-Perceived Ability, Attitude, and WTC

This section reports the findings on self-perceived ability in and attitude toward CL listening and speaking as well as WTC at Time 1 and Time 2 in students studying different modules. Table 1 shows the means and standard deviations of students' responses. Three 2 (time) \times 3 (group/module) mixed ANOVAs were conducted with self-perceived ability, attitude, and WTC as the respective dependent variables.

For self-perceived ability, there was a significant main effect of time, $F(1, 396) = 11.486, p = .001, \text{partial } \eta^2 = .028$. This suggested that, disregarding the module taken, students rated their CL listening and speaking ability significantly higher at Time 2 than at Time 1. A significant main effect of group was also found, $F(2, 396) = 11.858, p < .001, \text{partial } \eta^2 = .057$. Subsequent pairwise comparisons showed that, disregarding time, the Bridging students' self-rated listening and speaking ability was significantly lower than that of the Core students and the Enrichment students (both $ps < .001$). Although the ability rating of the Core group appeared lower than that of the Enrichment group, the difference was not statistically significant ($p = .291$). There was no significant time \times group interaction effect, $F(2, 396) = .896, p = .409, \text{partial } \eta^2 = .005$.

For attitude toward CL listening and speaking, there was a significant main effect of time, $F(1, 396) = 7.267, p = .007, \text{partial } \eta^2 = .018$. Disregarding module, students' attitude increased from Time 1 to Time 2. No significant main effect, however, was found of group, $F(2, 396) = 1.753, p = .175, \text{partial } \eta^2 = .009$, which

Table 1 Students' self-perceived ability in and attitude toward CL listening and speaking as well as willingness to communicate in CL

	Bridging ($N = 35$)		Core ($N = 113$)		Enrichment ($N = 251$)	
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Ability	3.25 (1.18)	3.56 (1.12)	3.65 (1.16)	4.06 (.82)	3.89 (.93)	4.09 (.84)
Attitude	3.64 (1.24)	3.85 (.92)	3.65 (1.12)	3.90 (1.10)	3.72 (1.04)	4.09 (.93)
WTC	2.86 (1.03)	3.16 (1.21)	3.05 (.82)	3.16 (1.10)	3.08 (.77)	3.23 (1.05)

indicated that, disregarding time, there was no significant module/group difference in children's attitude. There was no significant interaction effect between time and module, $F(2, 396) = .330, p = .719, \text{partial } \eta^2 = .002$.

Finally, for WTC, a similar pattern was found. A significant main effect was found of time, $F(1, 396) = 4.166, p = .042, \text{partial } \eta^2 = .010$. Not considering module, students' WTC significantly improved over a year. There was, however, no significant main effect of group, $F(2, 396) = .809, p = .446, \text{partial } \eta^2 = .004$. This suggested that, disregarding time, there was no significant difference in WTC in students studying different modules. There was no significant interaction effect, $F(2, 396) = .297, p = .743, \text{partial } \eta^2 = .001$.

6.2 Teacher Questionnaire

Tables 2 and 3 show teachers' perceptions toward the MC and Chinese characters, respectively. We compared how the perceptions may differ between more- and less-experienced teachers based on their total experience of CL teaching (0–5 years of teaching CL vs. more than 5 years) as well as experience of teaching under the MC (2 years or less vs. 3 years or more).³

As shown in Table 2, disregarding teaching experience, CL teachers overall had positive perceptions of oral language in CL teaching and learning ($M = 4.05$,

Table 2 Teachers' perceptions of the MC

MC perceptions	Total $M(SD)$	Teaching experience		MC experience	
		Less ($N = 177$)	More ($N = 134$)	Less ($N = 131$)	More ($N = 180$)
Oral language	4.06 (.370)	4.07 (.345)	4.05 (.400)	4.07 (.334)	4.05 (.395)
Accommodation	3.53 (.661)	3.48 (.595)	3.58 (.736)	3.50 (.579)	3.54 (.715)
Efficacy	3.90 (.359)	3.84 (.355)	3.97 (.351)	3.82 (.324)	3.96 (.373)

Table 3 Teachers' perceptions of Chinese characters

Chinese characters	Total $M(SD)$	Teaching experience		MC experience	
		Less ($N = 177$)	More ($N = 134$)	Less ($N = 131$)	More ($N = 180$)
Cultural values	4.28 (.432)	4.27 (.446)	4.27 (.415)	4.30 (.440)	4.26 (.427)
Importance of writing	3.90 (.538)	3.86 (.523)	3.94 (.558)	3.90 (.522)	3.89 (.551)
Student interest	2.67 (.736)	2.62 (.726)	2.72 (.747)	2.61 (.704)	2.72 (.757)

³Teachers were collapsed into these two broad groups based on teaching experience because a large majority of them, as mentioned earlier in the Teacher Questionnaire section, were in the early years of CL teaching (0–5 years: 56.9%) and the number of teachers for each of the other ranges of teaching experience was very small.

$SD = .370$). Their perceptions of the capability of the MC for accommodating different CL abilities and needs in students also appeared positive, but the rating was not particularly high on average ($M = 3.53$, $SD = .661$). Teachers' efficacy for using instructional strategies of differentiation to motivate and support student learning also appeared high ($M = 3.90$, $SD = .359$).

A set of independent samples t -tests was conducted to compare more- and less-experienced teachers. No significant difference was found for the total experience of CL teaching in terms of the perceptions of oral language ($t = .314$, $p = .754$) as well as those of the accommodating capability of the MC ($t = -1.381$, $p = .168$). This, however, was not the case for teachers' efficacy for differentiation in the classroom ($t = -3.247$, $p = .001$). Specifically, those more experienced in CL teaching ($M = 3.98$, $SD = .352$) were more efficacious than those who were less experienced ($M = 3.85$, $SD = .355$) in terms of using strategies of differentiation to motivate and support students from different language backgrounds or with different CL abilities. A similar pattern was found when the experience of teaching under the MC was the independent variable. No significant difference was found between those who differed in MC experience for the perceptions of oral language ($t = .424$, $p = .672$) as well as those of the accommodating capability of the MC ($t = -.544$, $p = .587$). For teachers' efficacy for differentiation in the classroom, however, a significant difference was found ($t = -3.336$, $p = .001$). Those more experienced with the MC ($M = 3.96$, $SD = .373$) were more efficacious than those who were less experienced ($M = 3.82$, $SD = .324$).

Table 3 shows the descriptive statistics of teachers' responses pertaining to perceptions of Chinese characters in CL teaching and learning. Overall, CL teachers had positive perceptions of the cultural values of Chinese characters ($M = 4.28$, $SD = .432$), believing that they are an important aspect of Chinese culture, which is a fundamental goal of CL education in Singapore (that is, learning CL for maintenance of ethnic and cultural heritage). They also, overall, believed that character writing is important in CL learning (e.g., character recognition/reading and composition) ($M = 3.90$, $SD = .538$). They, however, perceived student interest in writing characters to be low ($M = 2.67$, $SD = .736$).

Independent samples t -tests were conducted to compare more- and less-experienced teachers. Significant difference was found for none of the three variables. When the independent variable was general CL teaching experience, there was no significant difference in perceptions of the cultural values of Chinese characters, $t = .081$, $p = .936$; the role of character writing in CL learning, $t = -1.334$, $p = .183$; or student interest in writing characters, $t = -1.176$, $p = .240$. The same pattern was found when MC teaching experience was the independent variable. For the perception of the cultural values of Chinese characters, $t = .737$, $p = .462$; for the role of character writing, $t = .132$, $p = .895$; and for perceived interest in students in writing characters, $t = -1.278$, $p = .202$.

6.3 *Teacher Focus Groups*

This section reports some qualitative findings on how teachers perceived the MC regarding major recommendations by the CLCPRC (e.g., the modular approach, emphasis on oral language, and the “*Reading First, Write Later*” principle for Chinese characters), as well as major difficulties and challenges they experienced under the MC. Because of space limitations, we are not able to include any excerpts from the FGs to illustrate teachers’ views and discussions.

6.3.1 **General Perceptions of the MC and Its Underlying Considerations**

Overall, teachers commented positively on the MC in light of its module-based, differentiated approach. They, for example, shared that the MC drew their attention to the reality that students came from different language backgrounds and increased teachers’ awareness of how teaching could and should be differentiated to cater to different needs in students. They also commented that students’ learning interest (which was a backbone of the report of the CLCPRC and a key consideration of the MC) had noticeably increased. They noted that students in lower grades (before P3; see, however, the following section on difficulties and challenges) particularly enjoyed CL learning more than students under the old curriculum. Many teachers cited the deemphasis of frequent formal exams in P1 and P2 under the MC as enabling them to focus on student-centered activities such as role play, group demonstration, and “show and tell” presentation where students had opportunity to use CL for oral presentation and interaction. Those activities, according to teachers, were fun and liked by young children in P1 and P2 and boosted their interest in CL learning and use (some teachers, though, added that those methods of teaching can be very time-consuming). Some teachers further positively commented that the differentiated approach provided an opportunity for using English as a tool to scaffold early learning for those children who had had little CL exposure prior to primary school.

Nevertheless, teachers also expressed some concerns over the effectiveness of the MC, or the lack thereof, on actual learning outcomes in comparison to the old curriculum. Although the MC underscored oral communication, in the teachers’ view, students’ actual oral proficiency, compared to that of students they taught under the old curriculum, did not necessarily improve. In other words, although the MC was recognized to have achieved the planned goal in terms of boosting students’ CL learning interest (and the student survey showed student’ self-perceived competence in listening and speaking improved over a year of studying), the MC was not necessarily better than the old curriculum in boosting students’ CL proficiency. It is interesting to infer from this finding that teachers seemed to care much about students’ *actual proficiency* as evidence of effectiveness of curriculum and teaching, even though the primary consideration of the curriculum reform was for *learning interest*.

Another concern across all FGs was around the “*Recognize First, Write Later*” principle emphasized in the MC. While teachers agreed that the reduced number of characters for writing released some burden of memorizing stroke orders, some argued that the prescribed distinction between *renduzi* (characters for recognition) and *xiyongzi* (characters for writing) in textbooks was too artificial to represent what students would need for reading and writing purposes. A teacher, for example, was concerned that students sometimes wanted to write about ideas with words in their oral vocabulary but did not know how to write the characters for those words. Because the characters for intended use were not supposed to be a target for writing (i.e., not in the list of *xiyongzi*), she had to tell students to use alternative words with characters in the list. This was cited as limiting students’ writing potential and CL learning. Some teachers quoted pressure from parents as a reason for their reducing character writing in P1 and P2, despite the fact that, as will be discussed in detail below, those teachers knew very well that this instructional choice would make it very difficult for students to catch up from P3 when there was much emphasis on writing tasks (e.g., paragraph writing and composition where character writing is fundamental) required of all students.

6.3.2 Perceived Difficulties and Challenges

Teachers also reported a number of difficulties and challenges they had experienced in teaching under the MC. A concern reported repeatedly across all FGs was the lack of articulation between the curriculum for P1 and P2 and that for P3 onward. Although the teachers, as reported earlier, agreed that the emphasis on oral communication (in early grades) made CL learning interesting for students, particularly lower-ability students studying the Bridging module, the lack of attention to character writing in P1 and P2, which manifested the principle of “*Recognize First, Write Later*,” created a huge “gap,” “sudden transition,” and “leap forward” (words repeatedly used by teachers to describe the lack of connection) when students moved on to P3 where writing tasks were an essential component of the curriculum. Almost all participants in every FG group reflected on this issue and expressed worries and frustrations about students’ inability to deal with the sudden increase in writing tasks from P3, where composition started to bring too much strain on students. Teachers reported that deemphasizing character writing or written language in general in P1 and P2 caused a number of problems in subsequent years of learning, such as lacking in ability to write characters (using *pinyin*, the alphabetic system for annotating characters for recognition and early reading purposes, to replace a character would be considered an error), written language being too colloquial, and decreased interest in CL learning from P3 (despite the notable presence of interest in P1 and P2). Some, more specifically, noted that student workbooks from P3 barely included any listening/oral language practice (an emphasis in P1 and P2). Consequently, some schools, through their school-based curriculum, incorporated writing from as early as possible and introduced paragraph writing from P2 (as opposed to delaying it to P3).

Another lack of articulation discussed by some teachers was between CL education in early childhood/preschool and primary school. While the MOE oversees primary schools, early childhood education was overseen by the then Ministry of Community Development, Youth and Sports (now the Ministry of Social and Family Development). Teachers reported diverse CL abilities in P1 students due to the different educational experiences they had in childcare centers. While the MC considered students from different home language backgrounds, diversities in language background intertwining with those in preschool experience made the differentiation in P1 and P2 much more difficult. Some teachers reported that it was difficult for their schools, as a result, to properly place children into different modules. For example, among those from English-speaking homes, some had learned *pinyin* systematically in preschool whereas others knew very little; nevertheless, all were required to learn *pinyin* for about 10 weeks at the beginning of P1 to enable their learning to read (e.g., character recognition and *pinyin*-supported early textual reading).

Additionally, some teachers expressed that the Bridging module, with its designed purpose to help EC children from English-speaking families, was still very challenging for non-EC children (e.g., immigrants from non-Chinese-speaking countries in Southeast Asia) who typically had no exposure to CL at all prior to primary school but often studied the Bridging module in the same classroom with EC children who, despite having English as the predominant home language, usually had had some CL exposure before entering primary school. Those demographic and linguistic diversities together with diversities in CL exposure created additional challenges that, according to teachers, cannot be effectively addressed through the MC and made instructional differentiation much more complex and difficult. Teachers wished that additional support would be available to them and those students to “bridge” the gap of learning in the Bridging module. Some reported that their schools, as a result, did not strictly follow the module-based approach but emphasized distinguishing different CL levels within a class and using extra time outside CL classes to provide individual or small-group instruction for those who needed support and to make up for the components of the Bridging module those students missed.

As a result of the issue of lack of articulation, teachers reported that they were balancing between the requirements of the MC (emphasis on learning interest and oral language) and parents’ perceptions of children’s actual proficiency or learning outcomes as reflected in test performance. Parents were reported to have a concern that their children did well in P1 and P2 but fared badly upon reaching P3 due to the sudden emphasis on writing. Parents wanted their children to have fun with CL learning but also wanted to see the learning “materialized” in good test results. Teachers were under pressure to strike a balance between making learning interesting for students, especially those studying the Bridging module, in the early grades (P1 and P2) through various oral language activities and students’ decreased interest and drop in grades or test results in later grades (P3 onward).

CL is a high-stakes school subject in Singapore, as students’ results in the Primary School Leaving Examination (PSLE), which is taken in P6, have a

determinant effect on how they will be streamed to study different types of curriculum in secondary school. It is thus not surprising that any reform of CL education in Singapore would cause concerns not only in the general public but more saliently in teachers about student learning, classroom teaching, and test results. Some challenges and difficulties presented above already touched briefly on the issue of student assessment and CL testing and the dilemma teachers faced between student learning interest and test results. Regardless of the module studied, students all sit the PSLE, which is based on the Core module and has a heavy reliance on written language (e.g., passage comprehension and composition). This reality explains why across the FG groups, there were big concerns about how the MC created a challenge for preparing students for the PSLE (e.g., its lack of early attention to [character] writing), particularly those studying the Bridging module. While oral communicative skills are a key innovative point advocated in the new curriculum, it is the writing ability that determines the examination result. Teachers thus questioned the misalignment between the objectives of the MC and the reality of the examination, and wished that the testing system would be reformed to bring positive washback effects on classroom teaching.

7 Discussion

To answer the first research questions, student interest in CL (attitude toward CL listening and speaking and WTC in CL) increased over a year from P1 to P2 and so did their self-perceived ability in listening and speaking. To answer the second research question, teachers recognized the benefits of the differentiated approach of the MC for accommodating students with different CL abilities. The emphasis on oral language was underscored as boosting student interest in P1 and P2. Nonetheless, teachers were also concerned that the MC, compared to the old curriculum, did not seem to have enhanced students' oral proficiency, and the "*Recognize First, Write Later*" principle overall was not supported based on the many difficulties and challenges teachers experienced. Lastly, to answer the third research question, those difficulties and challenges included a sudden shift of the curriculum from a deemphasis on (character) writing in P1 and P2 to the strong inclusion of writing tasks and tests focused on written language from P3 onward. This lack of curriculum articulation or coherence, and the misalignment between the MC and the examination system, made teachers juggle curriculum requirements, parental expectations, students' difficulties and decreased interest from P3, and the need of written language skills for students to do well on high-stakes examinations. Consequently, teachers, who were on the front line of delivering the MC, had to be pragmatic and adaptive, exercising their agency to navigate many complex realities and act upon the MC in local and personalized ways.

In what follows, we discuss the findings in light of three complex realities, including (1) student interest versus CL proficiency; (2) oral language versus (character) writing and curriculum articulation; and (3) curriculum reform versus

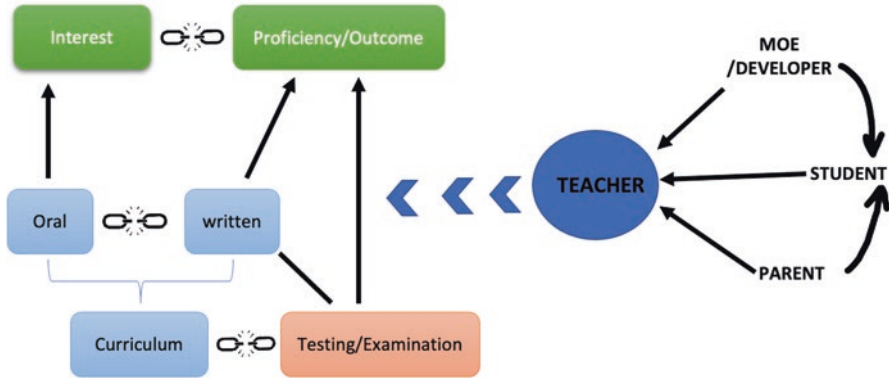


Fig. 1 Complex realities and conundrums faced by CL teachers

examination (see Fig. 1). Based on the findings (and more broadly sociolinguistics, bilingualism, and CL education in Singapore), we then discuss how curriculum innovation or reform is a collective enterprise and how boundary crossing is essential in curriculum evaluation toward evidence-based language-in-education planning.

7.1 *Student Interest Versus CL Proficiency as the Goal of Planning*

The first complex reality pertains to student interest versus CL proficiency as the goal of planning. While both should perhaps be a goal of language education or educational reform (given the literature on positive associations between learner interest and motivation and learning outcomes), the MC in CL education followed a local and pragmatic approach in that the MOE (policy maker and curriculum developer) prioritized boosting student interest over enhancing national CL achievements. The recommendations of the CLCPRC, and subsequently the MC, were based on pragmatic considerations of the changing realities of bilingualism in the country where it was considered unrealistic for everyone to be highly proficient in their MTL and equally proficient in English. This planning for “interest” encapsulated in the MC did seem to have achieved some effects.

In this respect, it is interesting, however, that teachers did not fully embrace the MC and were concerned that students’ actual proficiency was no better than that of students under the older curriculum. Under the consideration that students would not be able to have the (written language) skills to perform well on tests (see discussion later on examinations), schools and teachers customized approaches to boost learning outcomes, despite the recognition that student interest decreased as a result of those approaches. This choice of the teacher was influenced by parents and matched that of parents, who are also pragmatic stakeholders. Parents were happy

to see children have fun with CL learning (in P1 and P2) but quickly questioned teachers if the interest failed to be translated into good test results or outcomes.

The case of the MC may be unique in its “interest”-oriented planning for language education. Yet, the contentions revealed between different stakeholder groups (as a result of their different positioning and stakeholding interest), however, are by no means unique to Singapore. In fact, they have been widely reported in the literature on program innovation and evaluation and, more generally, language-in-education policy and planning (Donato & Tucker, 2010; Kiely & Rea-Dickins, 2005; Menken & García, 2010; Ross, 2009). Donato and Tucker (2010), in their evaluation of Japanese and Spanish FLES programs in the United States, for example, found that parents’ expectations, children’s views, and teachers’ views did not necessarily converge. Parents viewed the programs “as a vehicle for their children to develop cultural knowledge and awareness” as opposed to a certain level of “language proficiency per se” (p. 103). This might be attributed to their practical recognition that there was little opportunity or need outside the programs for children to use the target language. Yet, students themselves paid much attention to their language development and could use self-assessment tools to demonstrate abilities and identify areas they wanted to further develop. Teachers, likewise, emphasized skill attainment and proficiency development as important goals of student learning and had high expectations for bilingual language proficiency in students.

Although the FLES programs in Donato and Tucker (2010) and the MC in Singapore are arguably different in a number of dimensions, it is notable how much *teachers* in both contexts emphasized actual language learning outcomes in students. This is perhaps not a surprise given that teachers, in any context, are on the front line of program or curriculum delivery and thus reasonably aim to demonstrate effectiveness of teaching in terms of learner proficiency. The contrasting expectations of parents – proficiency and test performance in Singapore versus cultural knowledge and awareness in Donato and Tucker (2010) – seemed very reasonable in that CL in Singapore, as opposed to Spanish and Japanese in FLES in the United States, is a high-stakes school subject. This contrast is particularly noteworthy in a teacher’s remark during a FG that “They [parents] only look at grades or scores, caring little about how much their children’s cultural knowledge has expanded.” Such a highly pragmatic expectation of parents seems particularly interesting in that CL education in Singapore is primarily intended for cultural maintenance purposes (see also Chin, 2018).

7.2 Oral Language Versus (Character) Writing and Curriculum Articulation

The second complex reality that CL teachers negotiated and where their agency was demonstrated was about the lack of coherence in the MC. The MOE had a planned purpose to emphasize oral language in the MC. Students reported increased interest

in CL listening and speaking; parents, according to teachers, liked children's demonstrated interest in CL; and teachers were pleased to adopt student-centered activities to engage students in CL learning and use. These, however, were the case only in P1 and P2 where there was a deemphasis on formal testing and limited focus on character writing (and written language in general). From P3 onward, however, there was a sudden shift to strong written language requirements, and formal tests where written language was a heavy focus began to be introduced. As a result, according to teachers, students' interest decreased, their grades dropped, and parents complained (and had our longitudinal student survey carried on to later years beyond P2, we might see some change to the positive trend of increasing interest reported in this chapter). Some teachers/schools consequently chose not to strictly follow the principles underpinning the MC but started an emphasis on writing practice from as early as possible in P1 and P2 that, they believed, could bridge the gap and help make the transition from P1 and P2 to P3 smooth for students (and address their assessment needs; see discussion below on examinations) even though they were sometimes challenged by parents and were cognizant of the risk that this adaptation could demotivate students.

The conundrum induced by the lack of curriculum articulation and teachers' adaptation in curriculum enactment obviously have strong implications for the MOE, the primary user of the evaluation findings. These issues, however, are not unique to the current CL case. In fact, teachers' negotiation of curriculum requirements or policy mandates toward adaptive implementation, based on careful assessments of their local, micro contexts and realities of teaching (e.g., negotiating the interests and/or concerns of different stakeholder groups in the context of their own teaching or classroom realities), has been widely reported (e.g., Hyland & Wong, 2013; Ng & Boucher-Yip, 2016; Priestley et al., 2015). A gap is commonly seen between the intended/prescribed/planned curriculum and the enacted curriculum in language education (Menken & García, 2010; Orafi & Borg, 2009). The findings reported in this chapter, in this respect, have reaffirmed that it is crucial to consider teachers' perceptions, the micro contexts or local realities of their teaching, as well as the influence of those perceptions and negotiation of realities on actual teaching. For a program to achieve its planned goals (and to fine-tune toward sustainability), it is essential to engage teachers as key stakeholders or policy actors (Donato & Tucker, 2010; Kiely & Rea-Dickins, 2005; Menken & García, 2010).

7.3 Curriculum Innovation and Examinations

The last reality that concerned stakeholders, most importantly teachers, was the misalignment between some underlying principles of the MC and the examination system. In fact, many of the issues discussed earlier seemed to have a root in this reality. Teachers reported that CL tests (from P3 onward), particularly the high-stakes PSLE in P6, relied heavily on written language. This created a big gap between the planned goal of the MC that emphasized oral language proficiency and

learning interest and the immediate needs of students for achieving good test results with strong written language skills. CL teachers' enacted curriculum seemed to show a tendency, from the earliest years of primary school, to prepare students for high-stakes examinations as a result of negotiating the lack of coherent goals in the MC across learning stages and the misalignment between the MC and the examination system.

The gap between a reformed curriculum and an untouched gate-keeping examination system does not pertain to the CL case in Singapore alone. In the context of curriculum or program innovation, particularly in high-stakes educational reform contexts, it is not uncommon that reform efforts are encapsulated in curriculum contents and teaching materials (especially textbooks), without due attention to reforming high-stakes examinations and the impact of the gap on teachers and teaching (Agrawal, 2004). The influence or washback effect of testing on teachers and teaching is widely recognized and studied in language education (e.g., Cheng, 2005; Spratt, 2005). Language teachers are known to often "teach to the test" as a result of their passive and/or active goals for effective teaching, with instructional focuses and activities often guided by the content and format of a test or what they believe students need for gaining good test results (Cheng, 2005). They negotiate curriculum requirements and the interests of different stakeholder groups, including their own, which often results in a pragmatic approach of teaching with assessment goals strongly incorporated (Li & Baldauf, 2011). It thus seems no surprise that CL teachers reported adapting instructional focuses and rushing for full coverage of required curriculum content, even though they perceived learning gaps and decreased interest in students (from P3 onward) and consequently wished the goal planned in the MC for student interest, and the principles recommended by the CLCPRC, could transpire through a reformed examination system.

7.4 Curriculum Innovation and Evaluation, Language-in-Education Planning, and Boundary Crossing

Language education reform through program innovation is a collective enterprise where it is essential to engage diverse stakeholder groups or policy actors (e.g., curriculum developers, teachers, students, administrators, and communities), carefully analyze their different stakeholding interests, and evaluate how the diverse interests and positioning interplay to impact the goals planned of the innovation or reform (Donato & Tucker, 2010; Tucker, 2000a). Baldauf and colleagues (e.g., Baldauf et al., 2008; Kaplan & Baldauf, 1997) argued that language-in-education planning involves a number of objectives related to components or sub-systems of education: the target population, the teacher, syllabus, methods and materials, resources, assessment and evaluation. Baldauf et al. (2008) specifically listed eight processes or policy considerations, including access policy, personnel policy, curriculum

policy, method and materials policy, resource policy, community policy, evaluation policy, and teacher-led policy.

The case of the MC in CL education in Singapore suggested that these component policies, in program innovation or educational reform, are not always considered holistically and do not often operate in synergy as a result of the diverse agendas of different stakeholder groups. The primary agenda of the MOE was boosting and sustaining student interest for maintaining Chinese Singaporeans' ethnic and cultural identity, as opposed to boosting national achievements in CL (this made the MC case distinct from many others in the context of the standards-based movement, where policy decisions are often driven by neoliberal considerations). The parents' agenda was to see learning interest as well as concrete outcomes/test results in children. Students wanted to be engaged in learning; teachers wished to boost interest but had to be pragmatic so as to prioritize outcomes/test results over learning interest when the two were in conflict. These diverse stakeholding interests, together with the missing links in the reform effort, make program innovation and implementation very complex (see Fig. 1). They reflect the complex realities of changing sociolinguistic milieu, bilingualism and English-medium education, and the important gate-keeping role of high-stakes examinations in the educational system in Singapore.

To unravel the complexity and improve language education policy and practice, it is essential to cross boundaries between component policies and stakeholder groups and analyze their interplay through careful evaluations based on diverse sources and types of evidence. In this chapter, we have crossed methodological boundaries (e.g., qualitative and quantitative methods) and boundaries between stakeholder groups in our evaluation of the modular CL curriculum in Singapore. Although we had to limit our scope with a restricted focus on some findings on teachers and students, we hope the findings have achieved the purpose of exemplifying boundary crossing (e.g., stakeholder engagement and methodological pluralism) in language curriculum and program evaluation toward evidence-based language-in-education planning. Our approaches to boundary crossing have, in particular, underscored that evaluating a language program or curriculum is analogous to interpreting a "Necker Cube," to use Tucker's metaphor for describing language teaching (Tucker, 2000b, p. 26). A narrow attention to any single stakeholder or a narrow reference to any single source of evidence would obscure understandings about the complexity involved in curriculum reform and implementation and limit the generation and interpretation of evaluation findings.

8 Conclusion

We reported some findings, drawing upon student surveys and teacher FGs and surveys, of the evaluation of the MC, which was intended by the MOE, through a differentiated approach, to cater to different language backgrounds and CL abilities in primary school students in Singapore such that everyone could be supported to

achieve their best learning potential and develop a lasting interest in learning and using CL for maintaining their ethnic and cultural identity. We showed how key stakeholders, including teachers and students as well as parents and the MOE, bring into the process of implementing the MC considerations and perspectives which were not necessarily in synergy. This consequently resulted in teachers' pragmatic and adaptive approaches to curriculum/policy interpretation and implementation. We also demonstrated the importance of crossing boundaries between stakeholders and methods in program evaluation.

One of the core issues of program evaluation is the utilization of evaluation findings. This issue was not a focus of this chapter. Nevertheless, it is important to touch on it briefly in this conclusion. At different stages of the project, workshops and tailored reports were provided to participating schools, and meetings and discussions were conducted with the CL team of the CPDD, MOE. A full project report was also subsequently submitted to the MOE. It is noted, however, that while the evaluation project was being launched, another committee had been formed by the MOE to review MTLs (Malay, Chinese, and Tamil) in Singapore (MOE, 2011). The Mother Tongue Languages Review Committee (MTLRC) (MOE, 2011) aimed to build on earlier reviews, including the report of the CLCPRC (2004), to review the evolving sociolinguistics of MTLs and provide recommendations for MTL education and reform. To some extent, this reflects the Singaporean society's fast-responding approach to important issues like education. The MTLRC report reiterated the fundamental issue which formed the backbone of the CLCPRC report (2004), that is, the reality and continuing trend of English becoming a predominant home language of Singaporean children and the reform of curriculum and teaching toward developing lasting interests in students for using their respective MTL. Important issues such as better alignment between curriculum and examination, which emerged in the MC evaluation project, received much attention in the MTLRC report, which subsequently influenced the 2015 Chinese Language Syllabus (Primary) (MOE, 2014) and the format of the PSLE. From 2017, for *Huawen* (Chinese), for example, listening and speaking increased to 35% of the total score of the PSLE; and written composition, which can be taken with dictionary assistance, decreased to 20%. For *Gaoji Huawen* (Higher Chinese), however, the PSLE is still a fully written test.

Acknowledgement This chapter is based on a project funded by the Office of Education Research, National Institute of Education (NIE) (#OER52-08ZSH). We are grateful to a number of teachers and students for their participation and school principals for their support. Throughout the project, we received support from the Chinese Language team of the Curriculum Planning and Development Division of the Singapore Ministry of Education (MOE). Although the project was commissioned by the MOE and funded by NIE/MOE, any views expressed in this chapter are ours and do not represent theirs.

References

- Agrawal, M. (2004). Curricular reform in schools: The importance of evaluation. *Journal of Curriculum Studies*, 36(3), 361–379.
- Baldauf, R. B., Jr., Li, M., & Zhao, S. (2008). Language acquisition management inside and outside the school. In B. Spolsky & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 233–250). Blackwell.
- Borg, S. (2006). *Teacher cognition and language education: Research and practice*. Continuum.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.
- Chin, C.-K. (2018). Chinese language curriculum in Singapore (1960–2000): From culture transmission to language application. In K.-C. Soh (Ed.), *Teaching Chinese language in Singapore* (pp. 3–24). Springer.
- CLCPRC. (2004). *Report of the Chinese Language Curriculum and Pedagogy Review Committee*. Retrieved from <https://silo.tips/download/report-of-the-chinese-language-curriculum-and-pedagogy-review-committee>
- Davis, L. (2013). Language assessment in program evaluation. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley.
- Davis, J. M., & McKay, T. H. (Eds.). (2018). *A guide to useful evaluation of language programs*. Georgetown University Press.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge.
- Goh, H.-H. (2017). *Mandarin competence of Chinese-English bilingual preschoolers: A corpus-based analysis of Singaporean children's speech*. Springer.
- Hiver, P., Al-Hoorie, A. H., & Mercer, S. (Eds.). (2020). *Student engagement in the language classroom*. Multilingual Matters.
- Hu, B. (2010). The challenges of Chinese: A preliminary study of UK learners' perceptions of difficulty. *The Language Learning Journal*, 38(1), 99–118.
- Hyland, K., & Wong, L. L. C. (Eds.). (2013). *Innovation and change in English language education*. Routledge.
- Kaplan, R. B., & Baldauf, R. B., Jr. (1997). *Language planning: From practice to theory*. Multilingual Matters.
- Kiely, R., & Rea-Dickins, P. (2005). *Program evaluation in language education*. Palgrave Macmillan.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Newbury House Publishers.
- Li, M., & Baldauf, R. B., Jr. (2011). Beyond the curriculum: A Chinese example of issues constraining effective English language teaching. *TESOL Quarterly*, 45(4), 793–803.
- Liu, Y., & Zhao, S. (2008). *Coding the transformation of Chinese pedagogical practices in Singapore primary schools: A study of experiment* (Technical report). Center for Research in Pedagogy and Practice, National Institute of Education.
- Liu, Y., Zhao, S., & Zhang, D. (2006). *Chinese language instruction in Singapore primary school classrooms: A study of pedagogical practice* (Technical report). Center for Research in Pedagogy and Practice, National Institute of Education.
- Lynch, B. K. (2003). *Language assessment and programme evaluation*. Edinburgh University Press.
- Menken, K., & García, O. (Eds.). (2010). *Negotiating language policies in schools: Educators as policymakers*. Routledge.
- Ministry of Education (MOE). (2006). *2007 Chinese language syllabus primary*. Curriculum Planning and Development Division, Ministry of Education.

- Ministry of Education (MOE). (2011). *Nurturing active learners and proficient users: 2010 Mother Tongue Languages Review Committee report*. Ministry of Education.
- Ministry of Education (MOE). (2014). *2015 Chinese language syllabus primary*. Curriculum Planning and Development Division, Ministry of Education.
- Nation, I. S. P., & Macalister, J. (2020). *Language curriculum design* (2nd ed.). Routledge.
- Ng, P. C. L., & Boucher-Yip, E. F. (Eds.). (2016). *Teacher agency and policy response in English language teaching*. Routledge.
- Norris, J. M. (2006). The why (and how) of student learning outcomes assessment in college FL education. *Modern Language Journal*, 90(4), 590–597.
- Norris, J. M. (2016). Language program evaluation. *Modern Language Journal*, 100(s1), 169–189.
- Orafi, S. M. S., & Borg, S. (2009). Intentions and realities in implementing communicative curriculum reform. *System*, 37(2), 243–253.
- Pakir, A. (1992). Issues in second language curriculum development: Singapore, Malaysia, Brunei. *Annual Review of Applied Linguistics*, 13, 3–23.
- Priestley, M., Biesta, G. J. J., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury.
- Ross, S. J. (2009). Program evaluation. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 756–778). Blackwell.
- Spratt, M. (2005). Washback and the classroom: The implications for teaching and learning of studies of washback from exams. *Language Teaching Research*, 9(1), 5–29.
- Tucker, G. R. (2000a). Concluding thoughts: Applied linguistics at the juncture of millennia. *Annual Review of Applied Linguistics*, 20, 241–249.
- Tucker, G. R. (2000b). Precision, elegance and simplicity perspectives on TESOL and art. In C. Clyde (Ed.), *The journal of the imagination in language learning* (pp. 24–26). Center for the Imagination in Language Learning.
- Zhang, D. (2017). Word reading in L1 and L2 Chinese learners: Commonalities and differences. *Modern Language Journal*, 101(2), 391–411.
- Zhang, D., Zhao, S., & Li, L. (2014). Teachers' perceptions and use of information and communication technologies (ICT) in Chinese language education. In S. Li & P. Swanson (Eds.), *Engaging language learners through technology integration: Theory, applications, and outcomes* (pp. 239–258). IGI Global.
- Zhao, S., & Liu, Y. (2010). Chinese education in Singapore: Constraints of bilingual policy from the perspectives of status and prestige planning. *Language Problem and Language Planning*, 34, 236–258.
- Zhao, S. H., Liu, Y. B., & Hong, H. Q. (2007). Singaporean preschoolers' oral competence in Mandarin. *Language Policy*, 6, 73–94.

Impact of the COVID-19 Pandemic on American Students' Willingness to Study Abroad



Feng Xiao and Kun Nie

Abstract As an important extension of college education, study abroad has been consistently attracting American college students in the twenty-first century (for a review, see Institute of International Education, IIE open doors. Retrieved from <https://opendoorsdata.org/annual-release/u-s-study-abroad/?tab=us-study-abroad>, 2022), allowing students to cross geographic, cultural, and linguistic boundaries. However, the COVID-19 pandemic rerouted most American students back to the U.S. in spring 2020 and led to cancellations of most study abroad programs in the following summer and fall. The ongoing uncertainty of international travel and increasing concern about personal safety due to this pandemic has caused a decline in study abroad enrollments in the U.S. The pandemic has raised a need for reconceptualizing study abroad and reevaluating study abroad curricular requirements. The magnitude and duration of the impact of the pandemic on study abroad, including the current models that have been followed, cannot be properly evaluated without a nuanced understanding of students' willingness to study abroad in the coming years, their perceptions of the role of study abroad in their academic studies, and the impact of study abroad (or possible lack thereof) on their career prospects and life during a time full of uncertainties. This chapter used an online survey to examine 107 U.S. college students' willingness to study abroad and factors that can affect their decisions. Findings suggest that U.S. colleges need to cross boundaries between humanities/social sciences and STEM in curriculum design in order to meet the needs of the increased number of STEM majors who plan to study abroad after the pandemic. Moreover, the traditional Junior Year Abroad model (established in the early 1920s) needs to be modified in order to meet the emergent needs of freshmen and sophomores who are the majority of prospective study abroad students.

Keywords Study abroad · Higher education · Education policy

F. Xiao (✉) · K. Nie

Department of Asian Languages & Literatures, Pomona College, Claremont, CA, USA

e-mail: feng.xiao@pomona.edu; kun.nie@pomona.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_12

257

1 Introduction

Since the World Health Organization declared COVID-19 a pandemic on March 11th, 2020 (World Health Organization, 2020), education systems around the world have been heavily impacted by school closures. In the following month (April 2020), the number of students affected by school closures quickly reached 1.40 billion, the highest month figure from 2020 to 2021 (UNESCO, 2022). In the U.S., the average duration of school closures was more than 41 weeks from 2020 to 2021 (UNESCO, 2022). As the availability of vaccines significantly increased in 2021, most U.S. colleges switched back to in-person education in fall 2021, showing a high level of recovery in higher education. However, study abroad programs, which have suffered most from the pandemic, have shown little sign of recovery. For example, according to the Institute of International Education (IIE), there was a 53% year-to-year decline of U.S. students who studied abroad in the 2019–2020 academic year (IIE, 2022). In fact, some colleges have reconsidered the functionality of the study abroad office. For example, Pomona College (<https://www.pomona.edu>) changed the name of its study abroad office to the office of international and domestic programs in 2021, expanding the functionality of the original study abroad office due to the fact that most study abroad programs were still unavailable in 2021, and some only offered online international experiences. Given the persistence of the pandemic and complexity of COVID-19 elimination (Oliu-Barton et al., 2022), some scholars think that global youth mobility will be fundamentally altered by the pandemic, which might eventually lead to the end of all study abroad programs (Huish, 2021). The actual future of study abroad programs may not be so pessimistic, but policy makers in U.S. higher education need to act quickly in order to meet students' emergent needs for global experiences. To achieve this goal, it is indispensable for us to understand students' willingness to study abroad after the pandemic.

2 Defining the Concept of Study Abroad

The concept of study abroad often involves being temporarily outside one's home country for educational purposes. In practice, study abroad programs are highly diversified, varying in program design such as the setting, structure, curriculum, duration, and pedagogy (e.g., Giedt et al., 2015). As such, there are many definitions of study abroad (see, e.g., Collentine, 2009; Freed, 1995; Kinginger, 2009; Xiao, 2015). According to Xiao (2015), all these definitions share the following features: leaving one's home country for educational purposes, using the target language after class, and staying in a local community where the target language is widely used. These definitions emphasize settings where the target language is often the first language of people in the local community, whereas Glaser (2017) argues that the definition of study abroad needs to include settings where the target language

has the status of a lingua franca. For example, Stewart and Lowenthal's (2022) study focused on 15 international students who studied abroad in the Republic of Korea during the pandemic. The informants were from seven different countries, and English was used as a lingua franca while abroad (see also Wang & Diao, this volume). In order to include English or any other language as a lingua franca in academic settings, the present study defines study abroad as a pre-scheduled educational stay in a foreign country where learners can immerse themselves in the target language through communication with local people and/or through academic studies. This definition does not include virtual study abroad programs such as the virtual global internship program offered by the Council on International Educational Exchange (CIEE, 2022) because these online programs usually lack after-class social contact with local people from the target culture.

3 History of U.S. Students Studying Abroad

The origin of credit-bearing U.S. study abroad programs can be traced back to the early 1920s. In 1923, the Delaware Foreign Study Plan (also known as the Junior Year Abroad) was founded at the University of Delaware (Chieffo & Griffiths, 2004). In 1925, a group of 32 female students and a professor from Smith College completed a similar junior year abroad in France (Walton, 2005). Since then, the Junior Year Abroad model has been widely adopted by U.S. colleges. Most study abroad programs in the 1920s were based in West Europe. During the 1920s, IIE was the other agency that was actively involved in sending U.S. college students abroad for education (Hullihen, 1928). For example, in 1927, IIE officially started to grant scholarships to American undergraduate students who planned to study abroad (Hullihen, 1928). Since then, institutional efforts such as the Fulbright scholarship (starting in 1946), and the Boren awards (starting in 1994) have been increased to promote study abroad among U.S. college students.

In the past 30 years, there has been a swift increase in U.S. college students studying abroad. For every region of the world, the number of study abroad students has grown since 1999 (IIE, 2022). For example, Asia has seen a drastic increase in U.S. students studying abroad, from 8834 students in the 1999–2000 academic year to 40,602 students in the 2018–2019 academic year. To put this into historical context, the normalization of U.S.-China relations began with President Nixon's visit to Beijing in 1972 (Council on Foreign Relations, 2021). This coincided with a shift in U.S. students' interest in studying in Asia, meaning that U.S. foreign policy might be associated with countries where students are interested in studying. Similarly, Africa has also seen a substantial increase in U.S. students studying abroad, rising from 3368 students in the 1999–2000 academic year to 13,455 in the 2018–2019 academic year.

Originally, Western European countries were much more popular for American students to study abroad than Africa, Asia, and South America, but as time has gone on, these non-Western-European regions have grown in popularity among study

abroad students. In particular, regions such as East Asia, the Middle East, and Eastern Europe have become popular locations for American students to study critical languages. Critical languages are defined as “foreign languages that are critical to [the] national security and economic prosperity” of the United States (U.S. State Department, 2021). For example, a popular summer study abroad program is the Critical Language Scholarship (CLS), which is fully funded by the U.S. State Department. CLS sends American citizens abroad at no cost to study critical languages such as Chinese, Arabic, and Russian (U.S. State Department, 2021). CLS has played an important role in the increase of American students studying abroad in Africa and Asia.

In the Middle East and North Africa (MENA) region, there has been a massive increase in students studying in predominantly Arabic-speaking countries. In the 1999–2000 academic year, the majority of American study abroad participants in the MENA region studied in Israel, where the majority language is Hebrew and which has strong relations with the U.S. However, in the 2018–2019 academic year, there was a substantial increase in students in countries such as Jordan, the United Arab Emirates, and Morocco (IIE, 2022).

In terms of the demographics of American students studying abroad, there have been shifts in areas such as race and gender in the past 20 years. Since 2000, non-white students who study abroad have almost doubled, with an increase from 15.7% to 31.3% (IIE, 2022). With regard to gender, women have been the predominant gender group (IIE, 2022). Women accounted for 65% of the students studying abroad in the 1999–2000 academic year, and this proportion increased slightly to 67.3% in the 2019–2020 academic year. This indicates that women are more likely than men to study abroad. A possible explanation is that more women study subjects that allow or encourage study abroad. For example, there is a higher concentration of women who study humanities and social sciences (HSS) subjects in comparison to science, technology, engineering, and mathematics (STEM) subjects. Although women make up about half of the workforce in the United States, they only comprise 27% of the STEM workforce (U.S. Census Bureau, 2021). A larger percentage of students who study abroad major in HSS. The gender ratios of those fields seem to align closely with the gender ratio of study abroad participants. Therefore, our study reported in this chapter included a demographic question about gender in order to compare with the IIE Open Doors data.

4 Benefits of Study Abroad

There are two main benefits of study abroad: cultural learning and language learning.

According to Allen et al. (2006), study abroad creates the ideal space for cultural learning. Regardless of location, every study abroad experience will involve some level of cultural learning and development of intercultural competence. For example, even if an American student studies abroad in a linguistically similar place such as England or Australia, s/he will still benefit and learn from a different culture,

such as traditions and politeness, through communication with people from the local community because study abroad programs give students a space which can open opportunities for sharing culture and a variety of perspectives on information learned beyond the classroom setting (Davis & Coryell, 2019). The benefit of cultural learning is often tied with the use of the target language, meaning that students can develop their target language skills and cultural knowledge concomitantly. For example, Taguchi et al. (2016) found that intercultural competence (defined as cross-cultural adaptability) and social contact (defined as hours of using the target language in different situations) accounted for 37.7% of the target language proficiency gains over one semester of study abroad in China. Social contact mediated the relationship between intercultural competence and proficiency; in other words, cultural learning and target language acquisition can be connected in situations where learners engage in various communications with local people. These social contacts are unique to on-site programs. Virtual study abroad programs, in contrast, usually lack authentic situations where students can have meaningful communication as part of their daily routines. Experiencing a different culture in an authentic setting is of interest to most U.S. college students.

5 Impact of the COVID-19 Pandemic on Study Abroad

The COVID-19 pandemic has massively impacted the world and its political and social environments. What was once a world full of travel and cultural exchange rapidly became a world of insular, isolated communities. Although domestic travel, at least in the United States, continued to function throughout the pandemic, most international travel was shut down or at least highly restricted. In early March 2020, the United States issued a global travel health advisory encouraging all Americans traveling abroad to come back to the U.S. In particular, the U.S. Government restricted travel to the U.K., Ireland, and continental Europe, which affected international students studying abroad through American universities (Redden, 2020). By late March of 2020, American students studying abroad had been sent home by their programs or by request of the U.S. Government, therefore halting essentially all in-person study abroad programs.

The study abroad options for developing intercultural competence and target language skills during the COVID-19 pandemic were mostly halted due to American students' inability to travel abroad from 2020 to 2021. But this does not mean that all study abroad programs stopped during the pandemic. In fact, many programs decided to shift to an online curriculum in order to provide students with some semblance of foreign language and cultural learning. According to Svanholm (2020), 45.2% of over 7400 students surveyed indicated interest in a virtual study abroad program offered during the pandemic. Although virtual study abroad was popular during the pandemic (at the time of the present study, it had been the case), challenges associated with virtual study abroad programs such as lacking meaningful activities and a full cultural immersion experience (Liu & Shirley, 2021) may make

students switch back to traditional onsite study abroad when the COVID-related travel restrictions are loosened or removed.

The year 2021 saw an attempted return of in-person study abroad; however, it was a rough return. Due to the highly contagious Delta and Omicron variants of COVID-19, study abroad programs that were planned to happen in-person were moved online at the last minute. In fact, many study abroad programs running during spring and summer of 2021 were remote, meaning that participants experienced the entirety of the program through a computer screen. For example, Princeton University had to cancel at the last minute many of its study abroad programs for the spring 2022 semester. Programs in Czechia, Denmark, France, Ireland, Israel, Switzerland, and the United Kingdom were all cancelled due to the Omicron variant of COVID-19 (Elliott, 2021). Similar cancellations of study abroad programs can be found across U.S. colleges. On one hand, these cancellations demonstrate the vulnerability of study abroad programs due to the uncertainty of the pandemic. On the other hand, the cancellations have created difficulties for policy makers to estimate how fast study abroad programs can recover when the COVID-19 virus stabilizes because these cancellations make it hard to predict students' needs for study abroad and make it hard to provide suggestions for study abroad programs to meet these needs after the pandemic. No existing studies have directly examined these issues. To fill these gaps in the literature, the present study used an online survey to answer the following research questions:

1. What will the post-pandemic study abroad student body look like in terms of their background?
2. What factors can affect U.S. students' decisions to study abroad after the pandemic?
3. What changes to existing policies are needed to accelerate the recovery of study abroad programs?

6 Methodology

6.1 Participants

The participants were 107 American undergraduate students with an interest in study abroad after the pandemic. The participants were recruited through email solicitation, which explicitly stated that only those who were interested in study abroad after the pandemic were eligible to participate in this research. The participants came from six liberal arts colleges and five universities in the U.S. They were 69 females and 38 males. Their age ranged from 17 to 24, with an average of 19.

6.2 Instrument

The participants were asked to fill out an online survey that consisted of two sections. After signing a consent form online, they were asked to complete the survey. The first section of the survey asked questions about the participants' demographics. The second section consisted of 10 statements about the COVID-19 pandemic and study abroad, asking participants to rank their opinions on a Likert scale ranging from strongly disagree (1) to strongly agree (5). At the end of the survey, there was an optional question asking what other factors affect the decision to study abroad. This question gave the participants a space to further elaborate their decision-making process for study abroad and generated crucial qualitative data for a deeper analysis of the participants' thoughts and plans. It was beneficial to combine quantitative and qualitative data in this study because none of existing studies included factors associated with a public health crisis (e.g., a pandemic) as factors that can affect the decision to study abroad. The qualitative question (an open-ended question) was able to elicit factors that were not included in the preceding quantitative questions.

All the survey questions were developed in two phases. In phase one, factors that can affect the decision to study abroad were collected from previous studies (e.g., Bandyopadhyay & Bandyopadhyay, 2015; Liu et al., 2018). In phase two, six female and six male college students who studied abroad prior to the pandemic were interviewed with questions on factors identified in phase one to finalize the survey questions. In addition, questions related to the pandemic were created based on the prior literature on study abroad experience during the pandemic (e.g., Stewart & Lowenthal, 2022). The target factors were divided into two subcategories as shown in Table 1.

As shown in Table 1, there are four pandemic-related variables and four non-pandemic-related variables. The former set includes travel-related factors because a study abroad experience often starts with international travel. The latter set includes program-related factors because the program design directly affects the quality of study abroad.

The online survey was created using Google Forms. Participant recruitment was done through multiple rounds of email solicitation. The email explicitly stated that only currently enrolled U.S. college students who were interested in studying abroad after the pandemic were eligible for participating in this research. The solicitation email was sent to foreign language instructors at different U.S. colleges by the two authors. The contacted foreign language instructors were requested to forward the

Table 1 Factors affecting the decision to study abroad

Pandemic-related	Non-pandemic-related
1. Travel policies of the U.S.	5. Major and minor requirements
2. Travel policies of the target country	6. Scholarships/financial aid
3. Vaccine availability	7. A home-college affiliated program
4. Number of COVID-19 cases in the target country	8. A non-home-college affiliated program

solicitation email to their students and colleagues. The data collection lasted for 2 months in fall 2021, and two rounds of reminders were sent out by the two authors during the data collection period. Each participant was given a gift card of \$10 as compensation after they completed the survey. Originally, there were 112 submissions, but five of them were duplicates and were removed from data analysis. Participants' demographics, the average rating of each response to a Likert-scale question and its score range, and responses to the open-ended question were used to answer the three research questions.

7 Results

7.1 Demographic Background of the Study Abroad Students

Research question one asked about the demographic information about the possible post-pandemic study abroad cohort. In answering this question, we compared our findings with the Open Doors data (IIE, 2022) because the Open Doors data shows the most comprehensive information about U.S. college students who studied abroad prior to the pandemic.

Our results showed that more females were interested in study abroad after the pandemic. Sixty-four percent of the participants ($n = 68$) who completed the survey identified as female while only 36% identified as male ($n = 39$). The Open Doors data on gender showed a similar pattern, with approximately 67.3% of study abroad students being women in the 2019–2020 academic year. This means that the pandemic did not change the gender ratio of U.S. college students who were interested in study abroad.

Regarding the participants' majors, the data were divided into four categories: STEM (science, technology, engineering, and mathematics), HSS (humanities and social sciences), both, and undeclared. Table 2 shows the number of participants by the category of major.

The Open Doors data (IIE, 2022) shows that in the 2019–20 academic year, 68.2% of the US students who studied abroad were HSS majors and 24.5% of them were STEM majors, but our data showed that 43% of the participants were HSS majors and 40% were STEM majors. In other words, the participants' major types were split almost evenly between HSS and STEM. The discrepancy between our data and the Open Doors data is probably because the Open Doors data (IIE, 2022)

Table 2 Number of participants by the category of major

Major	Number of participants
STEM	43 (40%)
HSS	46 (43%)
Both	4 (4%)
Undeclared	14 (13%)

represented students who had studied abroad prior to their data collection time, whereas our data represented students who would plan to study abroad after our data collection time. These STEM students who were interested in study abroad (reported by our study) might not be able to do so because most study abroad programs are designed to specifically target HSS majors (e.g., Giedt et al., 2015). There was also a small number of participants (4%) who were studying both HSS and STEM subjects, such as doing a double major in Mathematics and Chinese. In addition, 13% had not declared a major at the time of data collection in fall 2021. These findings suggest that colleges and study abroad programs need to offer more STEM courses to attract STEM majors to study abroad after the pandemic.

With regard to the year at college, the findings showed that 57% ($n = 61$) of the participants were freshmen, 23% ($n = 25$) were sophomores, 12% ($n = 13$) were juniors, and 8% ($n = 9$) were seniors. In other words, more freshmen and sophomores planned to study abroad after the pandemic than juniors and seniors did. This may be because freshmen and sophomores were at their early stage of college studies and thus had more time to plan their study abroad. Juniors and seniors might not be equally optimistic about an opportunity to study abroad in the immediate future because the COVID-19 pandemic had caused many uncertainties for such an opportunity to be available.

7.2 Factors Affecting the Decision to Study Abroad

Research question two asked about factors that can affect U.S. students' decisions to study abroad. Table 3 shows the average ratings for all included factors.

In Table 3, Factors 1–4 are pandemic-related factors. Among them, the travel policy of the target country had the highest average rating (4.4 out of 5), whereas the

Table 3 Ratings for included factors

Factor	Average rating	Rating range
1. Travel policies of the U.S.	3.5	2–5
2. Travel policies of the target country	4.4	3–5
3. Vaccine availability	3.7	1–5
4. Number of COVID-19 cases in the target country	3.0	1–5
5. Major and minor requirements	4.3	3–5
6. Scholarship and financial aid	4.1	3–5
7. A home-college affiliated program	4.1	2–5
8. A non-home-college affiliated program	3.3	1–5

travel policy of the U.S. only had an average rating of 3.5. This might be because the COVID-related restrictions in the U.S. were relatively less strict compared to those in other countries. For example, Germany categorized the U.S. as a COVID high-risk area in January 2022, meaning that study abroad participants from the U.S. must complete a strict quarantine period (German Missions in the United States, 2022). Similarly, Canada required a 10-day quarantine for travelers (Government of Canada, 2022). Therefore, some American students might prefer a program with fewer restrictions to maximize their study abroad experience. On the other hand, students who are more anxious about the pandemic might prefer studying abroad through a program with stricter COVID rules. Vaccine availability also had a high average rating (3.7 out of 5). This might be because the participants believed in the effectiveness of vaccines. It may also partially explain why the number of COVID-19 cases in the target country had the lowest average rating (3.0/5) because if one was fully vaccinated, they might be less worried about being infected and hospitalized. Correspondingly, the average rating of personal anxiety of the pandemic was relatively low (3.2 out of 5). With increased availability of vaccines, it seems reasonable that the COVID-related anxiety level could decrease and be less of a concern for planning study abroad.

In the category of non-pandemic-related factors, scholarships and financial aid had an average rating above 4.1 (out of 5), meaning that despite the impact of the pandemic, this factor was still perceived as very important by the participants. Therefore, a suggestion on how to attract more U.S. college students to study abroad is to offer more scholarships and financial aid. For studying abroad, students need to decide between a home-college affiliated program and a non-home-college affiliated program. Our findings suggest that students may prefer the former (4.1) over the latter (3.3). One possible reason is that students may have more trust in a home-college affiliated program because such a program may be perceived as part of the home college, which is more reliable than a non-home-college affiliated program, especially during a public health crisis. For example, not only is it much easier for students to study abroad through their college as they will be able to transfer credits, but it can be a lot more affordable because students may be more likely to receive financial aid from their home college while abroad.

In addition to these eight factors captured in the Likert-scale questions, one open-ended question was also included to give participants an opportunity to share other factors that can affect their decision to study abroad. Forty out of the 107 participants answered this question. The 40 responses were coded using the following categories: family-related, personal-interest-related, study-abroad-program-related, and target-country-related factors. A family-related factor was defined as a factor that was related to family issues, such as “any issues regarding my family, such as health, finances, etc.” A personal-interest-related factor was defined as a factor related to a person’s special interests such as “opportunities to generate income.” A study-abroad-program-related factor was defined as a factor related to any aspect of the study abroad-program, such as “courses offered at the study abroad institution and the ease of getting courses counted for my major.” A target-country-related factor

was defined as a factor related to any aspect of the target country, such as “the geopolitical climate (of the target country).”

The two authors coded the 40 responses separately and then discussed discrepancies to reach a consensus on coding. Eight out of the 40 responses were reiterations of factors included in the survey. For example, “the ability to re-enter the States” can be considered a travel policy of the U.S. (see Factor 1 in Table 3), and “the COVID restrictions in the country of interest” can be considered a travel policy of the target country (see Factor 2 in Table 3). Therefore, these responses were not counted as other factors. Among the remaining 32 responses, 11 were family-related, two were personal-interest related, 13 were study-abroad-program related, and six were target-country-related. “Parents’ opinions” were mentioned five times, the highest among the family-related factors. “Participation of friends” and “living arrangements while abroad” both were mentioned five times, the highest among the study-abroad-program related factors. “Environment of the target country” was mentioned four times, the highest among the target-country-related factors. None of these high-frequency factors were mentioned by more than five participants. This means that these factors may reflect personal situations in decision making, suggesting that when a student chooses an appropriate study abroad program, individualized concerns and needs should be addressed.

7.3 Policies to Accelerate Recovery of Study Abroad Programs

Research question three asked about policies that can accelerate recovery of study abroad programs. First, our findings suggested that study abroad programs and U.S. colleges need to adjust their advertisement methods post-pandemic in order to reach a wider array of students. For example, the findings showed that freshmen and sophomores were the majority (74%) of the participants who planned to study abroad after the pandemic. This does not necessarily mean that they planned to study abroad during their first or second year at college, but policy makers in higher education need to cross boundaries between the traditional Junior Year Abroad model and a more adaptive model that may appropriately meet the needs of freshmen and sophomores. For example, a series of workshops can be designed for these students in order to orient them to the study abroad setting and get them well prepared with knowledge of cross-cultural differences alongside a good target language proficiency. It is also possible to determine eligibility for studying abroad based on students’ target language proficiency, cross-cultural competence, and knowledge of the target culture rather than their year in college. In other words, some freshmen and sophomores can be qualified for studying abroad without waiting until their junior year as long as they meet all language- and culture-related criteria. These may help the enrollments of study abroad programs quickly bounce back to the level prior to the pandemic. Second, students majoring in a STEM subject had almost equal interest in study abroad in comparison to their HSS counterparts, showing a need to cross boundaries of subjects and tailor the traditional study

abroad curriculum to accommodate both HSS and STEM students. By creating more course flexibility in STEM major requirements as well as offering study abroad programs with STEM courses as an option, American colleges will be able to encourage more students to study abroad and be able to diversify the population of study abroad students.

Our findings also revealed that the participants preferred studying abroad through a home-college affiliated program rather than a non-home-college affiliated program, suggesting that programs affiliated with a U.S. college may attract more students from their U.S. affiliation. This means that at the beginning of the recovery of study abroad programs, those affiliated programs need to be prioritized to generate a speedy recovery. Most of the participants (70 out of 107) indicated that they planned to study abroad in Fall 2022. This means American colleges may need to get their affiliated programs in shipshape ahead of this possible turning point. This policy can maximize the use of educational resources for a possible quick rebound when the COVID-19 virus stabilizes. Moreover, scholarships and financial aid can affect students' decisions because program affordability affects the diversity of the study abroad student body. Therefore, existing study abroad programs can consider offering new scholarships and more financial aid to attract financially disadvantaged students. Colleges also need to provide sufficient consultations to their students so that personalized concerns can be addressed before departure. With regard to pandemic-related factors, the participants rated travel policies of the target country as a top concern. This means that colleges and study abroad programs need to provide appropriate pre-departure guidance to familiarize students with those travel restrictions and let them know what resources are available in case of emergency while abroad.

8 Conclusion

The present study examined the impact of the COVID-19 pandemic on American students' willingness to study abroad. Findings suggest that U.S. colleges and study abroad programs need to get ready for a possible enrollment rebound when the COVID-19 virus stabilizes. Existing programs need to diversify their course offerings by crossing boundaries between HSS and STEM. The Junior Year Abroad model needs to be modified to better prepare freshmen and sophomores for studying abroad. Student's public health concerns need to be addressed by incorporating pandemic-related travel policies into pre-departure training and establishing available resources for emergency (e.g., a public health crisis). This mixed-methods study crossed boundaries between traditional quantitative and qualitative research in order to provide a comprehensive understanding of the factors that can affect the decision to study abroad after the pandemic, because public-health-related factors were not directly discussed in previous study abroad research. Future study abroad research may adopt this approach to cross methodological boundaries especially when newly emerging situations are under investigation.

References

- Allen, H. W., Dristas, V., & Mills, N. A. (2006). Cultural learning outcomes and summer study abroad. In M. Mantero (Ed.), *Identity and second language learning: Culture, inquiry, and dialogic activity in educational contexts* (pp. 187–213). Information Age.
- Bandyopadhyay, S., & Bandyopadhyay, K. (2015). Factors influencing student participation in college study abroad programs. *Journal of International Education Research, 11*(2), 87–94.
- Chieffo, L., & Griffiths, L. (2004). Large-scale assessment of student attitudes after a short-term study abroad program. *Frontiers: The Interdisciplinary Journal of Study Abroad, 10*(1), 165–177.
- CIEE. (2022). *Virtual global internship-8 week*. Retrieved from <https://www.ciee.org/go-abroad/college-study-abroad/programs/online/virtual-global-internship-8-week>
- Collentine, J. (2009). Study abroad research: Findings, implications, and future directions. In C. Doughty & M. Long (Eds.), *The handbook of language teaching* (pp. 218–233). Blackwell.
- Council on Foreign Relations. (2021). *Timeline: U.S. relations with China 1949–2021*. Council on Foreign Relations. Retrieved from <https://www.cfr.org/timeline/us-relations-china>
- Davis, B., & Coryell, J. E. (2019). Relationships, participation, and characteristics of a community of practice in short-term adult study abroad. *Adult Learning, 31*(1), 17–26.
- Elliott, J. (2021, December 17). *Most study abroad programs canceled for spring 2022*. Retrieved from <https://www.dailyprincetonian.com/article/2021/12/princeton-study-abroad-canceled-spring-2022>
- Freed, B. F. (1995). Second language acquisition in a study abroad context. John Benjamins.
- German Missions in the United States. (2022, January 25). *Coronavirus (COVID-19)*. Retrieved from <https://www.germany.info/us-en/covid-19/2321562>
- Giedt, T., Gokcek, G., & Ghosh, J. (2015). International education in the 21st century: The importance of faculty in developing study abroad research opportunities. *Frontiers: The Interdisciplinary Journal of Study Abroad, 26*(1), 167–186.
- Glaser, K. (2017). Metapragmatic perceptions in native language vs. lingua franca settings: Does target language status during study abroad make a difference? *Study Abroad Research in Second Language Acquisition and International Education, 2*(1), 107–131.
- Government of Canada. (2022). *Covid-19 global travel advisory*. Retrieved from <https://travel.gc.ca/travel-covid/travel-restrictions/isolation>
- Huish, R. (2021). Global citizenship amid COVID-19: Why climate change and a pandemic spell the end of international experiential learning. *Canadian Journal of Development Studies/Revue canadienne d'études du développement, 42*(4), 441–458.
- Hullihen, W. (1928). Present status of the “junior year abroad”. *The French Review, 1*(2), 25–37.
- Institute of International Education (IIE). (2022). *IIE open doors*. Retrieved from <https://opendoorsdata.org/annual-release/u-s-study-abroad/?tab=us-study-abroad>
- Kinginger, C. (2009). *Language learning and study abroad: A critical reading of research*. Palgrave Macmillan.
- Liu, Y., & Shirley, T. (2021). Without crossing a border: Exploring the impact of shifting study abroad online on students' learning and intercultural competence development during the COVID-19 pandemic. *Online Learning, 25*(1), 182–194.
- Liu, Y., Kamnuansilpa, P., & Hirofumi, A. (2018). Factors affecting international students' decisions on destination for studying abroad: A case study in China. *Frontiers of Education in China, 13*(1), 93–118.
- Oliu-Barton, M., Pradelski, B. S., Algan, Y., Baker, M. G., Binagwaho, A., Dore, G. J., et al. (2022). Elimination versus mitigation of SARS-CoV-2 in the presence of effective vaccines. *The Lancet Global Health, 10*(1), e142–e147.
- Redden, E. (2020, March 20). *COVID-19 disrupts international student exchange in both directions*. Retrieved from <https://www.insidehighered.com/news/2020/03/20/covid-19-disrupts-international-student-exchange-both-directions>

- Stewart, W. H., & Lowenthal, P. R. (2022). Distance education under duress: A case study of exchange students' experience with online learning during the COVID-19 pandemic in the Republic of Korea. *Journal of Research on Technology in Education*, 54(Suppl. 1), 273–287.
- Svanholm, A. G. (2020, April). *The impact of COVID-19 on study abroad: April 2020 survey results*. Retrieved from <https://institutions.educations.com/insights/student-survey-covid-19-and-study-abroad>
- Taguchi, N., Xiao, F., & Li, S. (2016). Assessment of study abroad outcomes in Chinese as a second language: Gains in cross-cultural adaptability, language contact and proficiency. *Intercultural Education*, 27(6), 600–614.
- U.S. Census Bureau. (2021, January 26). *Women are nearly half of U.S. workforce but only 27% of STEM workers*. Retrieved from <https://www.census.gov/library/stories/2021/01/women-making-gains-in-stem-occupations-but-still-underrepresented.html>
- U.S. State Department. (2021). *About the program. Critical language scholarship program*. Retrieved from <https://clscholarship.org/about>
- UNESCO. (2022). *Education: From disruption to recovery*. Retrieved from <https://en.unesco.org/covid19/educationresponse>
- Walton, W. (2005). Internationalism and the junior year abroad: American students in France in the 1920s and 1930s. *Diplomatic History*, 29(2), 255–278.
- World Health Organization. (2020, March 11). *WHO director-general's opening remarks at the media briefing on COVID-19*. Retrieved from <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19%2D%2D11-march-2020>
- Xiao, F. (2015). Adult second language learners' pragmatic development in the study-abroad context: A review. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 25(1), 132–149.

Writing as a Design Art: Crossing Boundaries Between Disciplinarity and Rhetoricity in a University Business Program



Ryan T. Miller , Silvia Pessoa, and David Kaufer

Abstract While writing courses often include instruction in rhetorical aspects of writing (i.e., learning to write), business content courses often assign writing as a tool for learning and assessing content knowledge (i.e., writing to learn), with little attention to students' rhetorical understanding of genres. This leaves students with an incomplete understanding of disciplinary genres and at a disadvantage in the workplace. We argue for a *writing-as-design* conceptualization of writing, which crosses the boundary between learning to write and writing to learn by helping students develop an awareness of purpose and reader experience within a disciplinary content course. Using a case study of business plan writing in a business content course, we examine instructional materials, interviews, and student writing to illustrate how business courses can leverage a writing-as-design conceptualization to improve student writing.

Keywords Writing as design · Learning to write · Writing to learn · Business plan writing · Business writing instruction

R. T. Miller (✉)

Department of English, Kent State University, Kent, OH, USA
e-mail: rmill129@kent.edu

S. Pessoa

Department of English, Carnegie Mellon University Qatar, Doha, Qatar
e-mail: spessoa@cmu.edu

D. Kaufer

Department of English, Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: kauffer@andrew.cmu.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_13

1 Introduction

In university-level education, students are expected to write a variety of genres in their discipline of study, and students' ability to understand and produce these genres is a crucial part of becoming a full-fledged member of their discipline (Canagarajah, 2002; Duff, 2001; Johns, 1997). In the field of business, the ability to communicate effectively in writing is valued by employers and is a key to obtaining a job and advancing in one's career (Bacon & Anderson, 2004; Conrad & Newberry, 2011; Lentz, 2013; National Commission on Writing, 2004).

Although content courses in business programs often require students to write discipline-specific genres (Zhu, 2004a), the teaching of writing is often seen by business faculty as being outside of their responsibilities. Many disciplinary (content-area) faculty feel that writing instruction is the domain of the English department and that required first-year writing courses should "ensure college students can write when they get to us" (Downs, 2013, p. 50). As a result, many disciplinary faculty use writing mainly as a tool to assess content knowledge, with little attention to students' rhetorical understanding of genres (Annous & O'Day Nicolas, 2015; Zhu, 2004b), often leading to faculty frustration with students' writing (Downs, 2013). Such a focus on acquisition of disciplinary content knowledge through writing is often termed *writing to learn* in the writing pedagogy literature.

On the other hand, instruction in academic writing courses, such as first-year writing courses housed in English departments, typically focuses on developing students' awareness of audience, purpose, and other rhetorical exigencies with the goal of preparing students for future academic and professional writing demands. Such courses often include writing tasks that are thought to be generalizable across a wide range of writing situations that students may encounter in their academic and professional lives. This approach is often termed *learning to write* in the writing pedagogy literature. After completing these courses, students are expected to know how to write and to be able to transfer these skills to writing in their disciplinary content courses. However, there is little evidence that such transfer occurs (Smit, 2004; Wardle, 2007).

Thus, there exists a boundary between, on one hand, disciplinary content-area education and its focus on knowledge acquisition through *writing to learn* (and accompanying frustration with student writing), and, on the other hand, writing courses with a focus on rhetorical understanding through *learning to write* (and an unrealized assumption that the writing skills students develop will transfer to writing in their field of study or work). Previous research has suggested a number of obstacles that prevent the crossing of this boundary, including inconsistency of terminology across writing and disciplinary content courses, a lack of expertise in writing instruction among disciplinary faculty and in content-area writing demands among writing faculty, and a lack of explicit connection between the writing tasks students are taught to do in writing courses and those they are assigned in content-area courses (Nelms & Dively, 2007; Wardle, 2007).

We contend that the aforementioned perceptual boundary between disciplinary knowledge and disciplinary literacy (or between becoming a businessperson and writing/communicating like a businessperson) should be crossed in any business program. In this chapter, we argue for a rhetorically informed approach to writing instruction within disciplinary content courses. Specifically, we draw on concepts from the field of design to argue for what we call a *writing-as-design* conceptualization of disciplinary writing instruction. A writing-as-design approach goes beyond the typical boundaries of content learning by teaching students how to effectively communicate content while considering rhetorical purpose and reader experience. In this chapter, we present a case study of how the teaching of business plan writing in an introductory business course was naturally conceptualized from a writing-as-design perspective, yielding positive outcomes in student writing at an English-medium university in the Middle East. Our point of departure is another introductory business course that relied on a writing-to-learn model that conceptualized writing as a tool for acquiring disciplinary concepts.

In the following section, we review the literature on business writing and discuss the writing-to-learn and learning-to-write conceptualizations of disciplinary writing and how they relate to writing as design. We then contextualize our study by describing a case study of an introductory business course which implemented a writing-to-learn model. Lastly, we provide a rich description of a writing-as-design case study, showing evidence of its positive impact on student writing.

2 Literature Review

2.1 *Situating Business Writing in Disciplinary Writing Research*

Writing is essential in business, and business employers prefer employees who can effectively communicate in writing for business purposes (Conrad & Newberry, 2011; Lentz, 2013). For this reason, many business programs offer undergraduate and/or graduate courses in business communication, focusing on the oral and written skills that business students need for the real world (Bogert & Butt, 1996; Knight, 1999; Sharp & Brumberger, 2013). While disciplinary writing instruction may be common in business communication or technical writing courses, many business content courses also include writing assignments, though often with little explicit writing instruction (Canesco & Byrd, 1989; Zhu, 2004a). From interviews with business content faculty, Zhu (2004a) found that the faculty saw writing as a critical skill for students' future success and that this importance was reflected in business school policies regarding the use of writing assignments in business content courses. However, business content faculty expected students to transfer general academic writing skills from English courses to business writing, an expectation that has been found to be unrealistic (Hyland, 2002; Smit, 2004; Spack, 1997;

Wardle, 2007). Many business faculty saw writing instruction as falling exclusively under the purview of the English department, with business content faculty having little to no role in teaching writing.

However, writing instruction is necessary in business courses since without instruction business students may not understand the varying expectations of different business genres (Canesco & Byrd, 1989; Forman & Rymer, 1999; Zhu, 2004a, b). In addition, although many genres in business courses are similar to professional genres, students often have difficulty making this link without explicit instruction (Bacha, 2003; Bacha & Bahous, 2008; Campbell, 2002). Focused writing instruction has been found to improve student performance on business writing (e.g., Campbell et al., 1999; Pittenger et al., 2006); however, many business content faculty feel that they do not have the tools necessary to effectively teach writing (Plutsky & Wilson, 2001). This may be one of the reasons that most investigations of business writing have occurred in the context of business communication courses, rather than business content courses. Although many business communication and technical writing courses emphasize students' learning of rhetorical and linguistic features of business genres, this is often not the case in many business content courses. Instead, writing in many business courses is thought of as a tool for displaying acquisition of disciplinary concepts, a conceptualization known as writing to learn. This can be contrasted with other conceptualizations of writing, including learning to write and writing as design.

2.2 *Conceptualizations of Writing*

Outside the community of writing teachers and researchers, a dominant view posits that writing knowledge is of a piece with conceptual knowledge. That is, to write is to know that which one writes about. The syntax required to formulate the propositions of a subject matter carry the load of syntax required to communicate it. It is thought that knowledge and expertise, thus, are sufficient (not just necessary) for communicating it. Reciprocally, under such a view, the primary focus of evaluations of writing is to test the writer's acquisition of knowledge. This view of writing and writing education accounts for the behavior of so many content teachers who support student exploration of a subject matter and then assign a paper, oblivious to an art of "textual making" above and beyond acquisition of the subject matter concepts and terms of art. It accounts for the continued association of student writing and "blue book" responses, where teachers spot-check students' responses for accurate understandings of subject matter concepts.

In writing studies and writing pedagogy, this conceptualization of writing is known as *writing to learn*, or "writing as a means of acquiring information, understanding concepts, and appreciating significance" (Broadhead, 1999, p. 19). This contrasts with a *learning to write* conceptualization, which focuses on "acquiring the socially-mediated communication skills and genre knowledge appropriate to a specific discipline" (Broadhead, 1999, p. 19). Whereas writing to learn uses writing

as a tool to stimulate and assess the learning of content knowledge through the act of writing (Britton, 1982; Knipper & Duggan, 2006), learning to write is concerned with the act of writing itself, including composing skills necessary for formulating ideas into text in ways that comport with the text's social context (Bazerman, 2016). In the context of business education, many business faculty see the writing in business content courses as writing to learn and writing in English and technical writing courses as learning to write (Zhu, 2004a).

A large body of writing-to-learn research has shown that writing can substantially benefit the learning of content knowledge. Graham and Hebert's (2011) meta-analysis of research investigating effects of writing instruction on reading comprehension and learning found that students' comprehension of content knowledge increased through writing about that knowledge and that increased amounts of student was related with increased comprehension (see also a review by Klein & Boscolo, 2016).

Although research has found positive effects of a writing-to-learn approach, such an approach is insufficient in a number of ways. First, it may not be the most efficient way to learn content knowledge. Penrose (1992) found that although it may encourage learning to an extent, "writing may not be the best choice of learning activity when the goal is simply to gather factual information" (p. 476), and that "students who studied for a test retained more facts from their reading than students who wrote essays" (p. 488). Secondly, a focus only on content knowledge may omit opportunities for students to be socialized into their discipline by learning the communication norms and expectations of that discipline (Carter et al., 2007). Lastly, a focus on concept acquisition lacks contextualization. Cox et al. (2009, p. 74) state that writing in many business courses is reduced to filling in "formulas and templates," a common writing-to-learn method also used in some other disciplinary writing contexts (see, e.g., Burke et al., 2006). While templates may be tempting for faculty and for students due to their ease of use, they lack the "considerations of audience, word choice, tone, and many other concerns that a writer must negotiate" in order to effectively communicate information in real-world business contexts (Cox et al., 2009, p. 74). Similarly, Jablonski (1999) argues that a template-based "fill-in-the-blanks" approach to business plan writing is limiting because it does not consider the "rhetorical demands of specific situations" (p. 110).

On the other hand, a strict learning-to-write approach is common in writing courses such as the first-year writing courses that are required at many universities. These courses aim to develop writing skills that students will need in their subsequent academic work across the curriculum and in professional contexts. Although there is variation across institutions and even among individual instructors, these courses typically focus on developing various kinds of knowledge necessary for generalized academic writing, including rhetorical knowledge (by analyzing and composing a variety of texts), critical thinking skills (by evaluating, interpreting, and synthesizing information from sources), an understanding of writing as a process (by drafting, reviewing, collaborating, and revising), and knowledge of conventional language use, source citation, and genre formats common in academic writing (Dryer et al., 2014). These are important and do lay a solid foundation; yet,

the expectation that students will be able to transfer these generalized skills to the specifics of disciplinary content-area writing often goes unfulfilled (Smit, 2004). Academic writing must be embedded in a particular context, or as Wardle (2017, p. 30) writes, “there is no such thing as writing in general; writing is always in particular.”

In the present study, we cross the boundary between writing to learn and learning to write with a conceptualization of *writing as design*, which focuses on how disciplinary content is communicated within the rhetorical exigencies of a specific disciplinary genre.

2.3 *Writing as Design*

A conceptualization of writing as design extends the learning-to-write approach commonly found in first-year composition to disciplinary content classrooms. Writing as design means that a writer uses language in purposeful ways. The writer takes rhetorical purpose, audience, and disciplinary context into account, and chooses language to create a designed effect on the reader. This may seem commonsensical, but as Kaufer and Butler (2000) argued in their theoretical formulation of writing as a design art, it is not. In contrast to writing to learn, writing as design emphasizes elements of written craft that support, but stand outside, subject matter foundations, such as rhetorical purpose and audience.

Relationships between writing and design have been discussed for many years (e.g., Buchanan, 1985, 1995, 2001, 2007; Crilly et al., 2008; Friess, 2010; Frith, 2004; Hart-Davidson, 2007; Kaufer & Butler, 1996, 2000; Sharples, 1996, 1999; Sheridan, 2010; Wrigley et al., 2009). Buchanan (1995) states that design and writing are similar in that they both have a rhetorical dimension and both are activities of invention. Drawing on Lawson’s (1997) definition of design, Sharples (1999) says that writing is a “conscious and creative communication with, and through, materials to achieve a human effect” (p. 60). Sharples outlines a number of similarities between writing and design, such as having goals that are flexible and broad and not having a singular process for achieving those goals. In addition, design tasks (including writing tasks) often do not have a defined end state, and designers (and writers) often continually try to improve the final product.

Kaufer and Butler (1996, 2000) formalized the conceptualization of writing as design under their theory of representational composition. They point out that theories and instruction of writing in schools in the U.S. have largely been dominated by structural accounts, and in response to this, Kaufer and Butler propose that writing be described and taught as a type of information design, envisioning texts not as words and clauses that form sentences and other linguistic units, but rather as ideas forming into design elements. In this theory, writers choose words and phrases to represent rhetorical effects that the reader experiences, creating an interactivity between writer and reader. Expert writers use their (sometimes implicit) knowledge of these design elements in ways that create desired effects in the reader. For

example, an appeal to a reader's values can be cued using words such as *justice*, *fairness*, and *happiness*; innovation can be cued by *breakthrough*, *cutting-edge*, and *state-of-the-art*; a result from a chain of thought can be cued by *because*, *owing to the fact*, and *on the grounds that*. As a skilled writer makes choices that invoke these rhetorical cues, there is a cumulative effect over the course of a text that shapes a reader's understanding of the text.

The substantial difference when writing is thought of as a design art is illustrated by Herrington (1985), who compared students' perceptions of writing in two chemical engineering courses, one a lab-based course and one a design-based course. Students in the lab-based course described the major purpose of writing as displaying the writer's knowledge; however, in the design-based course, students described the purposes as convincing the reader and making the text easier for readers to understand. Thus, thinking of writing from a design standpoint may considerably impact students' awareness of audience and purpose for writing. Similarly, Maun and Myhill (2005) describe high school students' perceptions of design in their written work. Even in high school, students were aware of design-related choices, and the effects that these choices had on the reader, such as using the words *exclusive* and *hand-crafted* to make a product sound more appealing. Although students were conscious of design-based linguistic choices, Maun and Myhill did not relate this to any specific genre or discipline.

Although the conceptualization of writing as design has been well developed, its application in teaching has not been explored extensively. Some studies that have investigated the teaching of writing as design have done so in the context of rhetoric education (e.g., Kaufer & Butler, 2000) or have focused largely on visual-spatial aspects of texts, rather than design of textual-semantic aspects (e.g., Hocks, 2003). In recent years, research on classroom implementation of a writing as design approach has begun to emerge (e.g., Helberg et al., 2018; Land, 2022; Wetzel et al., 2021), though little research has focused on writing in disciplinary content courses. The potential benefits of this perspective for disciplinary work were suggested by Ballard and Koskela (2013) for work in engineering, although without any description of how a writing as design could be implemented in instruction. Given the very functional nature of a writing as design perspective, it could be especially useful in disciplines that have highly functional, goal-oriented genres, such as business.

In the present study, we aim to answer the question *How can business courses cross the boundary between learning to write and writing to learn by leveraging a writing as design conceptualization to improve students' writing?* In answering this question, we present a case study of a business plan assignment in an introductory business course that naturally used a writing-as-design perspective. We provide an in-depth qualitative understanding of what is involved in a writing-as-design approach, illustrated by its application in teaching business plan writing and the impact on the resulting student writing.

3 The Present Study

3.1 Context

The present study arose from a larger study of disciplinary literacy development at an English-medium branch campus of an American university in the Middle East. Students at this campus are linguistically and culturally diverse, hailing mainly from the Gulf region, the greater Middle East, India, Pakistan, and Bangladesh. Seventy-five percent of students attended English-medium (or Arabic and English-medium) secondary schools, and the average TOEFL iBT score was 97, indicating that students' English proficiency was "good" or "high" (Educational Testing Service, 2015). The business program offers courses of the same standard as on the university's main campus in the U.S., and is accredited by the Association to Advance Collegiate Schools of Business (AACSB).

The larger study provided rare naturalistic data needed to understand teaching methods and student output in undergraduate education. We interviewed professors about the reading and writing demands of their courses, observed classes, and collected course materials and the writing that students produced. In our interviews with professors and classroom observations in the business program, we were particularly struck by one professor, Professor Parker,¹ who taught an introductory business course in which, without any formal training in rhetoric or business communication, he taught a sophisticated genre-based and rhetorical understanding of business plan writing which comported with the writing-as-design approach described earlier. This contrasted with some other business faculty, such as Professor Smith, who had a view of writing that was closer to writing to learn. In the present study, we describe how writing was approached in the context of the business plan assignments in Professor Parker's and Professor Smith's introductory business courses.

Both Professor Parker's and Professor Smith's courses were introductory business courses, and had the same objectives and used the same textbook.² Both Professor Parker and Professor Smith had extensive experience in industry prior to assuming their academic positions. Professor Smith was relatively newer to teaching, with 2 years of experience teaching only at the undergraduate level, while Professor Parker had 8 years of experience teaching both undergraduate and graduate courses.

¹All names are pseudonyms.

²Both courses were Introduction to Business; they had different course numbers because one (Professor Smith's) was for business majors and one (Professor Parker's) was for non-majors.

3.2 *Data*

In our analysis, we rely on 1-h audio-recorded interviews with each professor about the writing demands and expectations of their courses and their approaches to teaching business writing. In these interviews, we specifically asked the professors questions about the role of writing in their courses and what genres students were expected to write. In both courses, the main writing assignment was a business plan. Our questions focused on the professors' expectations for this genre, challenges students experienced, and the course materials used to scaffold students' writing of business plans (e.g., sample texts). We also examined the assignment guidelines each professor gave to students. Lastly, our data included the writing that students produced. In both courses, students completed the business plan assignment in groups of four to five students. Although previous studies have investigated business writing at a macro level by analyzing instructional practices (e.g., Albi et al., 2014; Jameson, 2006) or at a micro level by analyzing student writing (e.g., Lim, 2006), few studies have included both. Thus, our study also crosses a methodological boundary by including a variety of data sources in our analysis, including interviews with faculty, instructional materials, and student outcomes, allowing us to gain a more complete picture of a writing-as-design approach in a business content course.

Because the focus of this paper is on the writing-as-design approach, we provide an in-depth description of Professor Parker's approach to the business plan assignment and how this conceptualization of writing impacted student writing. For additional context, we also give a brief description of Professor Smith's approach. Our analysis of the business plans written by students included four business plans from Business 100 (taught by Professor Smith) and five business plans from Business 110 (taught by Professor Parker). In particular, we focused on the executive summary of each business plan since it was present in all of the texts and served as the orientation to the text.

3.3 *Text Analysis*

We conducted two types of text analysis, both of which reflect key rhetorical and linguistic aspects of the business plan genre. As the literature indicates, business plans are fundamentally arguments that use deductive logic and are generally based around a problem-solution genre structure (Lagerwerf & Bossers, 2002; Trailer & Wolford, 2001; Zhu, 2004b). Thus, conducting a genre analysis of the executive summaries seemed fitting. According to Swales (1990), genre analysis involves identifying the rhetorical moves and strategies in a text, and how these moves fit together to achieve a communicative purpose. Our analysis focused on the moves in the executive summaries and examined the ways these moves contributed to the expected problem-solution genre structure. Since business plans are meant to be arguments, and arguments tend to include a great deal of evaluation (Hood, 2010;

Liu & Thompson, 2009), our second type of analysis focused on the use of evaluations using the Appraisal framework from Systemic Functional Linguistics (Martin & White, 2005). Appraisal is a discourse analytic framework for the analysis of evaluations, and has been used extensively to investigate argumentation, stance taking, and reader positioning (see, e.g., Chang & Schleppegrell, 2011; Lancaster, 2014; Miller et al., 2014; Pessoa et al., 2017). This analysis allowed us to identify what was being evaluated, how it was evaluated, and how these evaluations contributed to meeting the problem-solution genre expectations. Together, these two analyses allowed us to provide a more complete picture of how a writing-as-design perspective can positively impact and enhance student writing.

4 Results and Discussion

We begin by briefly describing the business plan assignment taught by Professor Smith in Business 100, which reflected a writing-to-learn approach. We follow this with a more in-depth description of the business plan assignment taught by Professor Parker in Business 110, which reflected a writing-as-design conceptualization.

4.1 Business 100

Business 100 was an introductory business course taught by Professor Smith, who viewed writing as a tool for acquiring disciplinary knowledge. Similar to descriptions of business writing assignments by Cox et al. (2009) and Jablonski (1999), students in Business 100 were given a “Business Plan Template” and a sample business plan (based on the template). Students then wrote their business plans by filling in the template using concepts and terms of art from the field of business.

The business plan template consisted of sections, with a series of questions within each section. Each question had a “hint” describing how students should answer the question, with most (27 out of 37) targeting textbook knowledge, such as the following:

What form of business ownership (sole proprietorship, partnership, or corporation) will your business take? Why did you choose this form? *Hint: For more information on the types of business ownership, refer to the discussion of the different forms of business ownership in Chapter 3.*

All businesses have to deal with ethical issues. One way to address these issues is to create a code of ethics. List three core (unchanging) principles that your business will follow. *Hint: To help you consider the ethical issues that your business might face, refer to the discussion in Chapter 2.*

Provide a brief mission statement for your business. *Hint: Refer to the discussion of mission statements in Chapter 5. Be sure to include the name of your business, how you will stand out from your competition, and why a customer will buy from you.*

The business plan assignment description in Business 100 instructed students to “answer the questions that make up each part in the order that they are shown,” allowing a single possible order of information in the business plan. The professor emphasized the terms of art of business, stating in an interview that, “[business] is a new language for them, and at the end of each unit I list out the vocabulary or new words for them.”

The sample business plan based on the template (4483 words) was titled simply “Sample Business Plan,” and introduced a business called The Friendly Café. It was clearly written for pedagogical purposes to teach the business plan template to students. The sample was printed in black and white, had few figures, and was written in an informal and personal style with short sentences that used personal language and an informal tone, as in the following:

Our business will be a partnership. My family members and I will run the café together. We are a close family, and we work well together. Financially, we could not open the café without each other. Each partner will serve as a manager in the café. My mother will be the floor manager, my aunt will be the grill manager, and I will handle the accounting for the café.
(Business 100 Sample Business Plan)

The structure of the sample business plan reflected the focus on concept acquisition, with subheadings that corresponded to chapter titles in the course textbook. Students’ writing followed the template’s structure, and largely focused on how well the business idea illustrated business concepts.

Evaluations in the students’ texts largely focused on the proposed company or its product, as seen in the below excerpts (evaluations are in italics and the target of the evaluation is underlined):

Welcome to [Middle East] Student Tutoring’s business plan. MEST is a *one of a kind tutoring company* in [the Middle East]. [...] All in all our company offers *excellent tutoring services* with *maximum quality assurance* together with *maximum convenience* for the customer.

It is this belief in quality and the *importance* of our mission that makes us proud to provide knowledge to those who seek it.

MEST hires university students who will drive to their students’ homes and tutor them in a range of subjects. Our *highly competent tutors* will fulfill a number of requirements in order to ensure their abilities. First of all, they must be currently *enrolled at one of [the] highly selective universities [in the city]*.

Here, we see only positive evaluations, showing only that the company is *one of a kind* and that the services are *excellent*, with *maximum quality assurance*, *maximum convenience*, etc. Although this may show the quality of the company and its products, it does not demonstrate the need for the company within a problem-solution rhetorical mode. Overall, over 90% of evaluations in the Business 100 business plans were positive.

Overall, it was clear that the focus of the business plan assignment was the acquisition of business concepts, and there had been little attention to the use of rhetorically appropriate language to achieve the goals of a business plan.

4.2 *Business 110: A Case of Writing as Design*

Unlike Business 100, Professor Parker's business plan assignment in Business 110 did not make use of a template. Instead, Professor Parker's description of the business plan assignment was largely focused on the purpose and audience of the text, and what kind of effect the text should have on the reader. This is clearly seen in the assignment guidelines Professor Parker gave to students, which described a business plan as "essentially an argument" where the business plan writers need to answer the question: "Why is the investor better off with his or her money invested in my business?" Professor Parker describes the focus of the business plan as being on "what the business plan should *do*" which is to obtain funding by "[defining] a significant opportunity or problem" and "[providing] a credible strategy" to solve it. According to Professor Parker, "You have a winner if you solve an important, valuable problem." Clearly, considering the purpose and audience for a business plan within a problem-solution rhetorical mode is important to Professor Parker. He further explained this in an interview:

Business plans are written for an investor audience, so whether or not you are raising money you are writing in a style which will appeal to those potential investors in your business so the first thing that has to be communicated is a sense of excitement about the opportunity you are presenting [...]. Then you have to convince the audience that the opportunity represents a potential for a significant investment and then that the team of people representing the idea are the right team to execute it. (Professor Parker, interview)

Throughout his assignment guidelines, Professor Parker emphasized the importance of considering the audience and the purpose of writing a business plan:

Writing the plan concentrates analysis, forces decisions, and improves thinking. [The] process brings key employees together for a realistic, consistent plan. [It] provides a model from which changes can be evaluated. [It] provides a document for outside advice and budget costs to be established. [It gives an] indication of seriousness. [And it is] simplified communication. (Professor Parker, assignment guidelines)

Worth noting is that Professor Parker emphasized the importance of having the business plan in writing. According to him, "All outside supporters will ask for it. [And it] aids in attracting people and institutional support" (Professor Parker, interview).

Rather than focusing on filling out a template and using terms of art from the textbook, Professor Parker offered tips on how to best write the business plan to achieve the desired effect on potential investors. He stated in his guidelines that the students should have an argument that emphasizes "desirable characteristics" and a new, original idea. Practical writing tips are given as well: "Construct the argument in bullets, be sure each is supportable. Get it all down. Then worry about the order."

Once the ideas were down on paper, the students were advised to come up with “a great story line” and to use it to “connect with the audience, build credibility, [and] get and hold attention.” At the same time, because investors are busy, the students were instructed to “write for a skimming reader using figures and bullets.” In Professor Parker’s conceptualization of business plan writing, there was a clear audience in mind for the business plan, and writing was described not as using terms of art in a prescribed template, but as a process of brainstorming ideas and then arranging them to grab the audience’s attention, connect with them, and have a strong, credible argument. These ideas reflect a desire to achieve a human effect on an audience, a key aspect of writing as design discussed by Sharples (1999) and Kaufer and Butler (1996, 2000).

In addition to the assignment guidelines, Professor Parker also gave students three sample business plans written by previous undergraduate and graduate students. The businesses proposed in these plans were ones that had actually been funded, lending authenticity to them as successful business plans, rather than texts written for pedagogical purposes only. The differences from the sample business plan used in Business 100 were striking. First, the sample business plans in Business 110 looked like authentic business documents in terms of design. They were much longer (average number of words: 9262); they were printed in color; they included a title page with the name of the business (e.g., ClearCount Medical Solutions) followed by “Business Plan.” The names of the team members and the faculty advisor were listed, followed by a contact e-mail address. An outline with the content of the plan and the page numbers followed on the next page. Subheadings, lists with numbering and bullet points, tables, figures, and diagrams were used throughout the plan. Second, although the Business 110 sample plans had many of the same subheadings, such as *Executive Summary* and *Marketing Analysis*, each was different in layout and in their order of presentation of the information. Clearly, the writers of these plans did not follow a template and had put thought into how the information was structured and presented. Third, a disclaimer at the bottom of the title page made this document look more real and authentic. The disclaimer read:

This document and its contents are the sole property of the ClearCount team. By reviewing this document, you agree to the confidentiality of the information contained herein.
(Business 110 sample business plan)

Although the plans were written for the purposes of a class, the disclaimer shows that they were also written to be read by a specific audience and with the purpose of seeking funding for an original idea.

The most striking difference between the two sets of business plans was the emphasis in Business 110 on problem solving through an original idea. This original idea was set up as a solution to an existing problem. As in the Business 100 plans, the importance of the product was described; however, the Business 110 plans took into consideration how the product or company solved a problem rather than just describing the product or company itself. They referenced data and existing conditions and problems (using sources) that the business ideas would solve. They were also written in a more formal tone than the Business 100 sample plan

with references to the company (rather than the people who will run the company), the problems in the medical field they seek to solve, and how the proposed product aims to solve them. In line with the conceptualization as writing as design, the sample business plans in Business 110 read like arguments with the desired purpose of convincing a group of investors to fund a company that would be successful by solving an important problem.

This conceptualization of business plan writing resulted in student business plans that were structured using a problem-solution mode, as discussed in the literature (e.g., Trailer & Wolford, 2001; Zhu, 2004b). For instance, (1) is an excerpt from a Business 110 student business plan that starts with the description of a global situation with relevance to a local situation that has led to a problem, namely the high cost of telecommunication for immigrants to the country, as seen in (2). Given this situation and problem, the proposed product sought to provide a solution, namely the development of a new telecommunication solution, stated in (3). The introduction and description of the proposed product was followed by an evaluation of the limitations of existing solutions, as seen in (4).

- (1) Opportunity: The world serves as home to approx. seven billion (7 Bn) or more people. Not all these people are on the same piece of archipelago floating around. Yet rapid globalization has created the need for a person on one end of the world to communicate, perhaps with someone on the other end.
- (2) Taking [this country] as example, immigrants here do not always have the comfort of spending hours talking to their loved ones back in their home country because of high international calling charges failing to take advantage of handy and comparatively inexpensive VOIP service.
- (3) Mission Statement: MEVoip has come up with the perfect solution that makes use of today's advanced technology and global communication modules, VOIP to come up with one easy, efficient, effective and inexpensive service to bring you vocally closer to anybody, anytime, anywhere in the world.
- (4) Available Solutions: There were cheaper yet relatively efficient modes of communication available, for people who were aware of other options like VOIP made use of these services, but this still did not help the laborers. However there were limitations to these as they required net connectivity, a computer, things that were obviously hindrance to portable options of communications.

This analysis shows how the students in Business 110 wrote business plans that responded to a problem or situation that had not been adequately addressed by existing products, thus generating a need for their proposed product. The business plans were argumentative, used a problem-solution rhetorical mode, and sought to create a specific effect in the reader (Kaufers & Butler, 1996, 2000; Sharples, 1999), namely for the investor audience to feel that the company will be successful because it fills a need in society. As a reader and audience of a business plan, an investor may be more persuaded to fund a business plan that aims to solve a problem rather than one that focuses on describing the greatness of a product without having an immediate impact on society. As discussed by Lagerwerf and Bossers (2002), in the real world, an absence of the proper genre features, in this case the problem-solution features that create the argumentative nature of business plans, could decrease the likelihood of an investment.

We conducted an Appraisal analysis of the executive summaries in the five business plans produced by students in Business 110. In line with the use of a problem-solution rhetorical mode, the Appraisal analysis revealed that the targets of evaluations included market conditions, a problem that needs to be solved, and competitors and their products, as observed in (5) (targets of evaluations are underlined, evaluations are in italics).

- (5) *WiFiU's services solve two major problems: (1) the lack of relatively cheap, high-bandwidth, well distributed internet access to the estimated 750,000 active internet users in the [the country]. According to research done by our business, on average it costs an active internet user about [\$600]/month for high-profile internet access through the currently available telecommunication companies. Clearly, this amount of money is not affordable by the middle class citizens in [this country] (2) Marketing solutions that target specific customers who are present within the proximity of a certain business is currently not available in [the country].* (Business 110, business plan 1)

The phrase *solve two major problems* is a positive evaluation of WiFiU's services. Following this, there are negative evaluations of existing services: internet access to the country's users is lacking in terms of value, bandwidth, and distribution, while currently available telecommunications companies are evaluated as costing a large and unaffordable amount of money. Following this is a negative evaluation (*currently not available*) that shows the market conditions for the company's services.

Overall, the Business 110 plans included 30% negative evaluations and 70% positive evaluations. This is in comparison to only 10% negative evaluations and 90% positive evaluations in Business 100. The larger proportion of negative evaluations in Business 110 represents evaluations of the current situation, establishing the market for this new company by showing the current problem or the limitations of existing solutions or competitors, as observed in (5) above and in (6), (7), and (8) below. These evaluations help to establish the problem that is to be solved, showing the need for the product.

- (6) *The high demand in the market is not being satisfied, the market lacks ethics, the provided services have mediocre quality and lack comprehensiveness [...].* (Business 110, business plan 2)
- (7) *We have noticed that [the campus], a growing community, lacks many essential places for students living on campus.* (Business 110, business plan 3)
- (8) *The rates imposed by the single telecommunication firm would snatch off a high percentage of their income.* (Business 110, business plan 4)

In (6), by negatively portraying the current market condition as not satisfying the high demand of the customers and as being unethical and "mediocre," the student writers express the need for their product. Similarly, in (7), the student writers negatively evaluate the current situation of student housing by emphasizing how it "lacks many essential places for students living on campus" to propose the implementation of a convenience store on campus. In (8), the existing telecommunication company is evaluated negatively through a description of its high service fees being a financial burden. In contrast to the overwhelmingly positive evaluations in the Business 100 student writing seen earlier, the Business 110 plans strategically used negative evaluations to make the case for the proposed company and the problem it would

solve in order to more effectively convince an investor audience to invest in the company.

5 Conclusions and Implications

In this study, we have introduced a conceptualization of business writing instruction that crosses the boundary between disciplinarity and rhetoricity – that is, between accurate display of disciplinary knowledge on one hand, and strategic use of language to create a designed reader experience on the other. We have argued that this writing-as-design approach can be effectively applied in business content courses to enhance students' writing of business genres. This approach to writing is starkly different from a writing-to-learn approach that is often used in business content courses. While conceptualizing writing as a type of design has been discussed previously, there has been little research on actual application in the classroom, particularly in the context of disciplinary writing.

As exemplified by our case study of Professor Smith's course, Business 100, the goal of the business plan assignment was for students to acquire disciplinary concepts and vocabulary through the use of a template. This is similar to some previous descriptions of tools for writing to learn in content courses (e.g., Burke et al., 2006). However, as argued by Jablonski (1999) and Cox et al. (2009), the use of templates in business writing can be limiting because templates expose students to neither the situatedness of real-life writing nor the rhetorical elements that writers must consider to craft texts that communicate information effectively. The instructional goals of Business 100 concentrated largely on students' acquisition of the terms of art of business, and instructional materials did not mention the purpose or audience of the business plan or the goals of what the text should *do*. Student writing reflected this orientation by not implementing a problem-solution rhetorical mode, with little attention to persuasion of the audience to invest in the business. The Business 100 business plans were not structured as an argument to convince an investor audience of the feasibility of the business venture. In essence, business plan writing in Business 100 was not seen as the writing of a functional genre, but rather an opportunity to demonstrate disciplinary knowledge.

In contrast, Business 110 used a writing-as-design approach, which involved crafting a text to have specific effects on the reader, realized through linguistic and rhetorical choices (Kaufer & Butler, 1996, 2000; Sharples, 1999). This draws on learning-to-write principles that are common in courses such as first-year composition, but extends these to the specific context of a disciplinary content course. In line with notions of writing as design, purpose and audience were made explicit to students in the instructional materials used in Business 110. As a result, these business plans met the genre expectations of the business plan as an argumentative genre. According to Trailer and Wolford (2001) and Zhu (2004b), business plans are fundamentally arguments that use deductive logic, and are generally based around a problem-solution rhetorical mode, as observed in the Business 110 plans. This

conceptualization of business plan writing is likely to achieve a greater persuasive impact because an investor is more likely to fund a business that can show its feasibility by filling a real-world need.

We recognize that learning disciplinary concepts may be a necessary focus of introductory courses; however, this can be done through other mediums such as multiple choice or short-answer tests. As Penrose (1992) suggested, when the goal is to assess content knowledge, writing may not always be the best assignment. An additional problem with using disciplinary genre writing assignments for concept acquisition is that it, at best, gives students an incomplete understanding and, at worst, an incorrect understanding of real world writing expectations (as emphasized by Jablonski, 1999), potentially furthering the dissatisfaction that many businesses feel about new employees' writing (Casner-Lotto & Barrington, 2006). This was seen, for example, in the simplistic and overly positive Business 100 business plan sample. Engaging in disciplinary writing goes beyond displaying acquired knowledge; rather, "it is through writing in a discipline that students learn the patterns of reasoning, the habits of mind, of those in the discipline" (Penrose, 1992, p. 491), as emphasized by Professor Parker in Business 110. Conceptualizing writing as design can help faculty to develop students' understanding of reasoning and habits of mind in business, even among novice writers in introductory courses, such as Business 110. If given adequate tools, students are able to meet disciplinary expectations (see also Maun & Myhill, 2005).

Overall, a writing-as-design perspective crosses the boundary between the learning-to-write approach of general academic writing courses and the writing-to-learn approach that is common in disciplinary content courses. It does so through instruction that makes the rhetorical expectations of disciplinary genres explicit for students. Such instruction, however, necessitates that business faculty be aware of rhetorical expectations of the genre and the discipline and be able to communicate these to students in a clear manner (as Professor Parker was). We recognize that not all faculty are as rhetorically aware as Professor Parker in this study, and that for many faculty, rhetorical expectations are mostly implicit. Thus, there is a clear opportunity for conceptual and curricular boundary crossing wherein writing faculty work together with disciplinary faculty to analyze disciplinary genres and identify their rhetorical and linguistic features (for a discussion of the benefits of such collaboration in business content courses, see Carnes et al., 2001; for examples of such collaborations, see Gomez-Laich et al., 2019; Mitchell & Pessoa, 2019; Pessoa et al., 2018, 2019, 2022). Explicit instruction can be conducted through genre-based pedagogy (Martin, 1992; Rose & Martin, 2012), particularly utilizing the Teaching/Learning Cycle (Martin & Rose, 2005; Rothery, 1994). Such collaborations will allow faculty in the disciplines to engage in instruction of disciplinary genres that makes genre expectations clear and better prepares students for the real world.

The present study has illustrated boundary crossing in multiple ways. Methodologically, it demonstrated the benefits of crossing the boundary between macro and micro-analytic approaches to the study of language education. By analyzing a variety of data including course descriptions, assignment descriptions, interviews with faculty, and student writing, we were able to obtain a more

comprehensive understanding of what a writing-as-design approach to disciplinary writing looks like. Importantly, the present study crosses the conceptual boundary between disciplinarity and rhetoricity, or the acquisition of discipline-specific knowledge and the ability to express that knowledge effectively in writing. This is particularly important for curriculum and instruction in English-as-medium-of-instruction (EMI) contexts, such as the present study, because students may need more explicit instruction on how to use language effectively for real-world purposes in their field of study.

References

- Albi, A. B., Natividad, J. V., Pilar, O. L., & Conde, T. (2014). A genre-based approach to the teaching of legal and business English: The GENTT specialized corpus in the LSP classroom. In E. Bárcena, T. Read, & J. Arús (Eds.), *Languages for specific purposes in the digital era* (pp. 177–196). Springer.
- Annous, S., & O’Day Nicolas, M. (2015). Academic territorial borders: A look at the writing ethos in business courses in an environment in which English is a foreign language. *Journal of Business and Technical Communication*, 39(1), 93–111.
- Bacha, N. N. (2003). English across the academic and professional communities: A study of EFL learners’ needs at the Lebanese American University. *The Official Journal of the Association of American International Colleges and Universities*, 2, 33–62.
- Bacha, N. N., & Bahous, R. (2008). Contrasting views of business students’ writing needs in an EFL environment. *English for Specific Purposes*, 27, 74–93.
- Bacon, D. R., & Anderson, E. S. (2004). Assessing and enhancing the basic writing skills of marketing students. *Business Communication Quarterly*, 67, 443–454.
- Ballard, G., & Koskela, L. J. (2013, August). *Rhetoric and design*. Paper presented at the International conference on engineering design, Seoul, Korea.
- Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 11–23). Guilford.
- Bogert, J., & Butt, D. (1996). Communication instruction in MBA programs: A survey of syllabi. *Business Communication Quarterly*, 59, 20–44.
- Britton, J. (1982). Writing to learn and learning to write. In G. M. Pradl (Ed.), *Prospect and retrospect: Selected essays of James Britton* (pp. 94–111). Boynton/Cook Publishers.
- Broadhead, G. J. (1999). Addressing multiple goals for engineering writing: The role of course-specific websites. *Language and Learning Across the Disciplines*, 3(2), 19–43.
- Buchanan, R. (1985). Declaration by design: Rhetoric, argument, and demonstration in design practice. *Design Issues*, 2(1), 4–22.
- Buchanan, R. (1995). *Rhetoric, humanism, and design*. University of Chicago Press.
- Buchanan, R. (2001). Design and the new rhetoric: Productive arts in the philosophy of culture. *Philosophy and Rhetoric*, 34, 183–206.
- Buchanan, R. (2007). Strategies of design research: Productive science and rhetorical inquiry. In R. Michel (Ed.), *Design research now: Essays and selected projects* (pp. 55–66). Birkhäuser Board of International Research in Design.
- Burke, K. E., Greenbowe, T. J., & Hand, B. M. (2006). Implementing the science writing heuristic in the chemistry laboratory. *Journal of Chemical Education*, 83, 1032–1038.
- Campbell, N. (2002). Getting rid of the yawn factor: Using a portfolio assignment to motivate students in a professional writing class. *Business Communication Quarterly*, 65, 42–51.

- Campbell, K. S., Brammer, C., & Ervin, N. (1999). Exploring how instruction in style affects writing quality. *Business Communication Quarterly*, 62(3), 71–86.
- Canagarajah, A. S. (2002). *A geopolitics of academic writing*. University of Pittsburgh Press.
- Canesco, G., & Byrd, P. (1989). Writing required in graduate courses in business administration. *TESOL Quarterly*, 23, 305–316.
- Carnes, L. W., Jennings, M. S., Vice, J. P., & Wiedmaier, C. (2001). The role of the business educator in a writing-across-the-curriculum program. *Journal of Education for Business*, 74, 216–219.
- Carter, M., Ferzli, M., & Wiebe, E. N. (2007). Writing to learn by learning to write in the disciplines. *Journal of Business and Technical Communication*, 21, 278–302.
- Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century US workforce*. Conference Board, Inc., Partnership for 21st Century Skills, Corporate Voices for Working Families, Society for Human Resource Management.
- Chang, P., & Schleppegrell, M. (2011). Taking an effective authorial stance in academic writing: Making the linguistic resources explicit for L2 writers in the social sciences. *Journal of English for Academic Purposes*, 10, 140–151.
- Conrad, D., & Newberry, R. (2011). 24 business communication skills: Attitudes of human resource managers versus business educators. *American Communication Journal*, 13, 4–23.
- Cox, M., Ortmeier-Hooper, C., & Tirabasi, K. E. (2009). Teaching writing for the “real world”: Community and workplace writing. *The English Journal*, 98, 72–80.
- Crilly, N., Good, D., Matravers, D., & Clarkson, P. J. (2008). Design as communication: Exploring the validity and utility of relating intention to interpretation. *Design Studies*, 29, 425–457.
- Downs, D. (2013). What is first-year composition? In R. Malenczyk (Ed.), *A rhetoric for writing program administrators* (pp. 50–63). Parlor Press.
- Dryer, D. B., Bowden, D., Brunk-Chavez, B., Harrington, S., Halbritter, B., & Yancey, K. B. (2014). Revising FYC outcomes for a multimodal, digitally composed world: The WPA outcomes statement for first-year composition (version 3.0). *WPA: Writing Program Administration*, 38, 129–143.
- Duff, P. A. (2001). Learning English for academic and occupational purposes. *TESOL Quarterly*, 35, 606–607.
- Educational Testing Service. (2015). *TOEFL for academic institutions: Interpret scores*. Retrieved May 1, 2015, from <http://www.ets.org/toefl/institutions/scores/interpret>
- Forman, J., & Rymer, J. (1999). Defining the genre of the “case write-up”. *Journal of Business Communication*, 36, 103–133.
- Friess, E. (2010). Designing from fata: Rhetorical appeals in support of design decisions. *Journal of Business and Technical Communication*, 24, 403–444.
- Frith, S. (2004). A primitive exchange: On rhetoric and architectural symbol. *Architectural Research Quarterly*, 8, 39–45.
- Gomez-Laich, M. P., Miller, R. T., & Pessoa, S. (2019). Scaffolding argumentative analytical writing in a design class: A corpus analysis of student writing. *Linguistics and Education*, 51, 20–30.
- Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review*, 81, 710–744.
- Hart-Davidson, B. (2007). Shaping texts that transform: Toward a rhetoric of objects, relationships, and views. In C. Lipson & M. Day (Eds.), *Technical communication and the world wide web* (pp. 27–42). Lawrence Erlbaum.
- Helberg, A., Poznahovska, M., Ishizaki, S., Kaufer, D., Werner, N., & Wetzel, D. (2018). Teaching textual awareness with DocuScope: Using corpus-driven tools and reflection to support students' written decision-making. *Assessing Writing*, 38, 40–45.
- Herrington, A. J. (1985). Writing in academic settings: A study of the contexts for writing in two college chemical engineering courses. *Research in the Teaching of English*, 19, 331–361.

- Hocks, M. E. (2003). Understanding visual rhetoric in digital writing environments. *College Composition and Communication*, 54, 629–656.
- Hood, S. (2010). *Appraising research: Evaluation in academic writing*. Palgrave Macmillan.
- Hyland, K. (2002). Specificity revisited: How far should we go now? *English for Specific Purposes*, 21, 385–395.
- Jablonski, M. (1999). Teaching the complexity of business proposals. *Business Communication Quarterly*, 62, 108–111.
- Jameson, D. (2006). Teaching graduate business students to write clearly about technical topics. *Business Communication Quarterly*, 69, 76–81.
- Johns, A. M. (1997). *Text, role, and context*. Cambridge University Press.
- Kaufert, D. S., & Butler, B. S. (1996). *Rhetoric and the arts of design*. Lawrence Erlbaum.
- Kaufert, D. S., & Butler, B. S. (2000). *Designing interactive worlds with words: Principles of writing as representational composition*. Lawrence Erlbaum.
- Klein, P. D., & Boscolo, P. (2016). Trends in research on writing as a learning activity. *Journal of Writing Research*, 7, 311–350.
- Knight, M. (1999). Writing and other communication standards in undergraduate business education: A study of current program requirements, practices, and trends. *Business Communication Quarterly*, 62, 10–28.
- Knipper, K. J., & Duggan, T. J. (2006). Writing to learn across the curriculum: Tools for comprehension in content area classes. *The Reading Teacher*, 59, 462–470.
- Lagerwerf, L., & Bossers, E. (2002). Assessing business proposals: Genre conventions and audience response in document design. *Journal of Business Communication*, 39, 437–460.
- Lancaster, Z. (2014). Exploring valued patterns of stance in upper-level student writing in the disciplines. *Written Communication*, 31, 27–57.
- Land, C. L. (2022). Recentring purpose and audience as part of a critical, humanizing approach to writing instruction. *Reading Research Quarterly*, 57(1), 37–58.
- Lawson, B. (1997). *How designers think: The design process demystified*. Architectural Press.
- Lentz, P. (2013). MBA students' workplace writing: Implications for business writing pedagogy and workplace practice. *Business Communication Quarterly*, 76, 474–490.
- Lim, J. M. H. (2006). Method sections of management research articles: A pedagogically motivated qualitative study. *English for Specific Purposes*, 25, 282–309.
- Liu, X., & Thompson, P. (2009). Attitude in students' argumentative writing: A contrastive perspective. *University of Reading Language Studies Working Papers*, 1, 3–15.
- Martin, J. R. (1992). *English text: System and structure*. John Benjamins.
- Martin, J. R., & Rose, D. (2005). Designing literacy pedagogy: Scaffolding democracy in the classroom. In J. Webster, C. M. I. M. Matthiesen, & R. Hasan (Eds.), *Continuing discourse on language: A functional perspective* (pp. 251–280). Continuum.
- Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Palgrave Macmillan.
- Maun, I., & Myhill, D. (2005). Text as design, writers as designers. *English in Education*, 39, 5–21.
- Miller, R. T., Mitchell, T. D., & Pessoa, S. (2014). Valued voices: Students' use of engagement in argumentative history writing. *Linguistics and Education*, 28, 107–120.
- Mitchell, T. D., & Pessoa, S. (2019). Scaffolding second language disciplinary writing in Qatar: A case study of a design teacher's development. In L. Seloni & S. Henderson Lee (Eds.), *Second language writing instruction in global contexts: English language teacher preparation and development* (pp. 150–172). Multilingual Matters.
- National Commission on Writing. (2004). *Writing: A ticket to work... or a ticket out: A survey of business leaders*. College Board.
- Nelms, G., & Dively, R. L. (2007). Perceived roadblocks to transferring knowledge from first-year composition to writing-intensive major courses: A pilot study. *WPA: Writing Program Administration*, 31, 214–240.
- Penrose, A. N. (1992). To write or not to write: Effects of task and task interpretation on learning through writing. *Written Communication*, 9, 465–500.

- Pessoa, S., Mitchell, T. D., & Miller, R. T. (2017). Emergent arguments: A functional approach to analyzing student challenges with the argument genre. *Journal of Second Language Writing, 38*, 42–55.
- Pessoa, S., Mitchell, T. D., & Miller, R. T. (2018). Scaffolding literacy at a branch campus of an American university in the Middle East: Interdisciplinary collaborations. In M. Rajakumar (Ed.), *Western higher education in global contexts* (pp. 133–147). Lexington Books.
- Pessoa, S., Gomez-Laich, M. P., Liginlal, D., & Mitchell, T. D. (2019). Scaffolding case analysis writing: A collaboration between IS and writing faculty. *Journal of Information Systems Education, 30*, 42–56.
- Pessoa, S., Mitchell, T. D., Gomez-Laich, M. P., Maune, M., & Le Roux, C. (2022). Scaffolding the case analysis in an organizational behavior course: Making analytical language explicit. *Journal of Management Education, 46*, 226–251.
- Pittenger, K. K. S., Miller, M. C., & Allison, J. (2006). Can we succeed in teaching business students to write effectively? *Business Communication Quarterly, 69*, 257–263.
- Plutsky, S., & Wilson, B. A. (2001). Writing across the curriculum in a college of business and economics. *Business Communication Quarterly, 64*, 26–41.
- Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School*. Equinox.
- Rothery, J. (1994). *Exploring literacy in school English (write it right: Resources for literacy and learning)*. Metropolitan East Disadvantaged Schools Program.
- Sharp, M. R., & Brumberger, E. R. (2013). Business communication curricula today: Revisiting the top 50 undergraduate business schools. *Business Communication Quarterly, 76*, 5–27.
- Sharples, M. (1996). An account of writing as creative design. In C. M. Levy & S. Ransdell (Eds.), *The science of writing* (pp. 127–148). Lawrence Erlbaum.
- Sharples, M. (1999). *How we write: Writing as creative design*. Routledge.
- Sheridan, D. M. (2010). Fabricating consent: Three-dimensional objects as rhetorical compositions. *Computers and Composition, 27*, 249–265.
- Smit, D. W. (2004). *The end of composition studies*. Southern Illinois University Press.
- Spack, R. (1997). The acquisition of academic literacy in a second language. *Written Communication, 14*, 3–62.
- Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.
- Trailer, J., & Wolford, C. (2001). On the logic of composition in writing business plans: Teaching argument structure. *New England Journal of Entrepreneurship, 4*, 41–50.
- Wardle, E. (2007). Understanding ‘transfer’ from FYC: Preliminary results of a longitudinal study. *WPA: Writing Program Administration, 31*, 65–85.
- Wardle, E. (2017). You can learn to write in general. In C. E. Ball & D. M. Loewe (Eds.), *Bad ideas about writing* (pp. 30–33). West Virginia University Libraries Digital Publishing Institute.
- Wetzel, D., Brown, D., Werner, N., Ishizaki, S., & Kaufer, D. (2021). Computer-assisted rhetorical analysis: Instructional design and formative assessment using DocuScope. *Journal of Writing Analytics, 5*, 292–323.
- Wrigley, C., Popovic, V., & Chamorro-Koc, M. (2009). A methodological approach to visceral hedonic rhetoric. In *Proceedings of the International Association of Societies of Design Research conference 2009* (pp. 1–11). The International Association of Societies of Design Research.
- Zhu, W. (2004a). Faculty views on the importance of writing, the nature of academic writing, and teaching and responding to writing in the disciplines. *Journal of Second Language Writing, 13*, 29–48.
- Zhu, W. (2004b). Writing in business courses: An analysis of assignment types, their characteristics, and required skills. *English for Specific Purposes, 23*, 11–135.

Part IV

Language-in-Education Planning and Policy

Introduction and the Tuckerian Impact

Around the globe, language planning in education or language-in-education planning has never been a peripheral issue in research on educational policy and policy-related debates. There are policies, *de facto* and *de jure*, that regulate whether a language should be taught, when it is to be taught, and what the medium of instruction should be, among many other issues. Language-in-education planning serves diverse agendas and purposes, whether educational, social, cultural, economic, and/or political. For example, the linguistic imperialism of the English language has resulted in the promulgation of English language education, including English-medium education, in many countries of the world, framed to improve their nation's access to scientific and technological advancement and to the global marketplace. Likewise, diverse societies, particularly post-colonial societies, in facing old and new challenges (e.g., immigration), are intrinsically and/or extrinsically motivated to adopt a bilingual/multilingual education policy, whether it is for promotion of social cohesion or maintenance of a heritage language. Needless to say, shifting geopolitics globally have also motivated competitive agendas—overt or covert—that shape policies on language in education.

These agendas on language education policy are concurrent with a standards-based movement in education in many places in the world where, driven sometimes by a neoliberal ideology on education and to prepare the nation for the knowledge economy of the 21st century, visions and goals are set for education to enforce accountability (e.g., the Common Core State Standards in the United States). For example, language standards with proficiency benchmarks are often outlined to guide the development of language curriculum, assessment, and teaching at critical stages of schooling and education (e.g., ACTFL, 2012; WIDA, 2014). High-stakes proficiency tests and assessment procedures are also often developed to serve a gate-keeping role or function as a hidden policy (see Shohamy, 2006). Likewise, teacher standards are developed to outline the dispositions that teachers need to adopt toward students and the profession and the knowledge and skills they need to

have for effectively teaching and assessing students, such as the *ACTFL/CAEP Program Standards for the Preparation of Foreign Language Teachers* (ACTFL/CAEP, 2013) and TESOL International Association's (2019) *Standards for Initial TESOL Pre-K-12 Teacher Preparation Programs*.

Understanding, interpreting, and evaluating language-in-education planning, or what Cooper (1989) called "acquisition planning," necessitates attention to policy actors (e.g., teachers, communities, students, and administrators), micro contextual factors (e.g., classroom realities and access to resources), and importantly, if not more so, the macro social, cultural, and political backdrop of the formulation and implementation of the policy. Kaplan, Baldauf and their colleagues argued that language-in-education planning involves a number of objectives related to the target population, teacher supply, syllabus, methods and materials, resources, assessment and evaluation (Baldauf, 2005; Kaplan & Baldauf, 1997, 2003). Specifically, Baldauf, Li, and Zhao (2008) listed eight processes or considerations about policy decisions in language-in-education planning. These include access policy, personnel policy, curriculum policy, methods and materials policy, resources policy, community policy, evaluation policy, and teacher-led policy. Without a "holistic" consideration of these component policies and their complex interplay, an understanding of language education policy and processes will be impaired. In other words, boundary crossing is needed for the micro and macro issues that underpin the other parts of this volume, including language learning and the personal and social "consequences" it brings (Part I); teachers and teaching (e.g., teacher quality, teacher development, and instructional processes; Part II); and program implementation and evaluation (e.g., shared vision and coordinated effort across stakeholders; Part III).

Engagement with policy and planning issues runs through Dick's more than 50 years of scholarship on language education. This is probably the most "visible" line of contribution that Dick has made to language education and applied linguistics and where he has crossed many boundaries to improve public awareness and policy formulation toward a language-competent society and educational equity. In fact, the considerations that form the aforementioned component policies in language education planning were spotlighted by Dick several decades ago in his vision and discussion on a language-competent American society (e.g., Tucker, 1984). Throughout his career, Dick was on the frontline of language education policy discussion and debates. Not only did he conduct cutting-edge research on innovative language education to contribute to policy formulation, but he wrote prolifically on policy issues to increase public awareness of language in society, language learning, and bilinguality. During his tenure at McGill University, Dick was a language specialist for the Ford foundation. In that role he engaged in various sociolinguistic surveys and policy-related projects in Asia and Africa. Later, he brought the assets from those projects and the St. Lambert project (e.g., the impacts of dual language instruction on schooling, education, and society) to his new role as the Director of the Center for Applied Linguistics, where he was engaged in wider policy-related discussion and debates. This was further carried on to his tenure at Carnegie Mellon

University, where he aspired to shape a policy on bilinguality (Tucker, 2004) and global literacy at the University and beyond (Nair et al., 2012).

In the United States, Dick's publications have been cited to inform policy and legislative decisions (e.g., Tucker, 1986, which was later included as additional material to inform the congressional subcommittee hearing on English Language Constitutional Amendments). Through a number of publications and together with his collaborators, Dick discussed, debated, and contested the English-only movement in the country and pointed out the negative psychological, social, and educational consequences of English-only policy on American society (e.g., Padilla et al., 1991). He contested the assimilationist, "immersion" approach to coping with educational challenges among linguistically and culturally diverse students. Dick was among the earliest who sharply pointed out that additive or "genuine" bilingualism should be the goal of educational policy and practice: language minority or linguistically and culturally diverse students need to be supported to develop strong proficiency in English for academic learning and educational success; their heritage languages are a national resource; and English-speaking language majority students should be given an opportunity to learn another language (e.g., Lambert & Tucker, 1981; Tucker, 1981, 1984, 1991, 1997).

Exploring the research-practice-policy nexus is a notable characteristic of Dick's boundary crossing in light of the focus of the present part. This is seen in Dick's long-term commitment to *evidence-based* policy discussion, debates, and engagement. Dick's insights into acquisition planning, and his powerful arguments and advocacy for language education and bilinguality, as outlined earlier in this part and throughout this volume, are solidly based on the empirical findings that he and his collaborators generated as a result of crossing a myriad of linguistic, methodological, institutional boundaries (Donato & Tucker, 2007, 2010; Tucker & Donato, 2003; see also the introductions of previous parts). For example, Dick's insights into, and efforts to promote, early foreign language learning were solidly based on his longitudinal research on child language development and bilingual schooling (Part I); classroom processes and teaching (Part II); and program implementation and evaluation (Part III). To illustrate, one of the "hot" issues in language education that has strong policy implications is the age of language learning/beginning instruction. Dick's research on young Japanese as a Foreign Language (JFL) learners in the FLES project (Donato et al., 1994, 1996; Tucker et al., 1996) purposefully explored this issue. The findings showed that, overall, the younger cohort (grades K-2) were not at all disadvantaged in their JFL learning and development compared to the older cohort (grades 3-5). The younger cohort actually showed a more positive orientation toward the JFL program and seemed more engaged with homework. Additionally, the younger cohort seemed to have progressed much faster over the three years of JFL learning in their receptive vocabulary growth than did the older cohort. While the younger students performed less well than the older students initially (at the end of the first year of learning), two years later (that is, after they had learned Japanese for three years), they actually outperformed the older students. The findings thus cast doubt on "the claim that older students ... are categorically better equipped to learn foreign languages in the classroom than younger students;"

and led to the conclusion that “as the profession discusses the inclusion of foreign language courses in the core elementary curriculum and the appropriate starting point for children ... ‘younger is better’ from the perspective of overall receptivity and positive affective reactions to language learning” (Donato et al., 1996, p. 524).

Also noteworthy examples of boundary crossing in Dick’s policy-related scholarship are his global perspectives on language education, bilingualism, and multilingualism (Tucker, 1998, 2001a; cf. Donato, 2013). In many ways, it was through these global perspectives that Dick has helped us understand how bilingualism/multilingualism, as opposed to monolingualism, is the norm; how bilinguality is a social asset; and how simple “borrowing” in acquisition planning and policy formulation without careful contextualization or contextual assessment should be questioned. In fact, Dick’s scholarship on policy has touched on many countries and societies. Other than the many lines of his work outlined earlier in the North American context, Dick, in collaboration with others, surveyed English language policy in Jordan (Harrison et al., 1975) and English teaching and teacher training in China (Cowan et al., 1979), and analyzed educational policy in the Philippines (Tucker, 1987). He outlined suggestions for two-way bilingual education for heritage language students in Puerto Rico (Tucker, 2005; see also Tucker, 2008). And together with former graduate students, who are also contributors to this volume, Dick surveyed Confucius Institutes in the United States in the context of the Chinese government’s global strategy to promulgate Chinese language and culture (Li & Tucker, 2013) and Chinese as a second language education policy in Hong Kong for immigrant students from South Asia (Ke & Tucker, 2016).

What is more noteworthy is Dick’s effort to cross boundaries between policy contexts and directly compare them to enlighten us on language education, bilinguality, and society. For example, Dick contrasted the pervasive monolingualism in the United States with how “in the Netherlands, it is not at all unusual for students to study English, and either German or French, if not both” (Tucker, 1997, p. 91) and powerfully argued that “the development and the nurturing of such multi-language proficiency is *never* a liability—it is *always* an asset” (Tucker, 1997, p. 97; italicized emphasis original). Likewise, he drew our attention to the complex interplay between linguistic, social, economic, and political factors in language education planning and bilingual education through comparative analysis of the situations in Nigeria, Sudan, and the Philippines where “critical attributes” varied such as “the role and status of the languages in question, the presence or absence of a rich literary tradition in the language(s) in question, the availability of a sufficient quantity of trained teachers who are mother tongue speakers of the particular language(s), parental and community expectations concerning the educational chances and choices available to the youngsters” (Tucker, 1981, p. 10). Based on “case” analysis and comparisons, Dick (Tucker, 1981; see also Tucker, 1984 as well as Dutcher’s [1995] report for the World Bank, for which Dick acted as a consultant) further argued that language minority children “should be introduced to schooling in their home language should they and their parents so desire.” Similarly, he emphasized that:

the Canadian research on immersion programs for language-majority children should not be indiscriminately generalized and interpreted as offering evidence to legitimize the submersion of language-minority youngsters in American schools. Rather, those data should only be used to support what I hope will become increasingly frequent inquiries by English speaking parents concerning the desirability of enrolling their language-majority youngsters in bilingual programs. (Tucker, 1981, p. 11)

He decried how the French immersion model and findings (Lambert & Tucker, 1972) were often miscited and misused for policy formulation in the US context without reference to the nuance of demographic, sociolinguistic, and socio-political distinctions (Tucker, 1986; see also McField, 2014). Likewise, when discussing age of beginning instruction in a foreign language, Dick (Tucker, 2001b), despite the support from the JFL program (e.g., Donato et al., 1996) for “younger is better,” cautioned that countries or contexts of planning for early language instruction vary in a number of micro and macro factors, and “the validity of the adage ‘earlier is better’ would seem to depend at least partially on the optimization” of those factors (p. 597). (See also the discussion in Part III on language program “transferability.”)

References

- ACTFL. (2012). *ACTFL proficiency guidelines 2012*. American Council on the Teaching of Foreign Languages.
- ACTFL/CAEP. (2013). *ACTFL/CAEP program standards for the preparation of foreign language teachers*. ACTFL/CAEP.
- Baldauf, R. B. Jr., Li, M., & Zhao, S. (2008). Language acquisition management inside and outside the school. In B. Spolsky & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 233–250). Blackwell.
- Baldauf, R. B., Jr. (2005). Language planning and policy research: An overview. In E. Hinkel (ed.), *Handbook of research in second language teaching and learning* (pp. 957–970). Lawrence Erlbaum.
- Cooper, R. L. (1989). *Language planning and social change*. Cambridge University Press.
- Cowan, J. R., Light, R. L., Mathews, B. E., & Tucker, G. R. (1979). English teaching in China: A recent survey. *TESOL Quarterly*, 13(4), 465–482.
- Donato, R. (2013). Tucker, G. Richard. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley–Blackwell.
- Donato, R., & Tucker, G. R. (2007). K-12 language learning and foreign language educational policy: A school-based perspective. *Modern Language Journal*, 91(2), 256–258.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.
- Donato, R., Antonek, J. L., & Tucker, G. R. (1994). A multiple perspectives analysis of a Japanese FLES program. *Foreign Language Annals*, 27(3), 365–378.
- Donato, R., Antonek, J. L., & Tucker, G. R. (1996). Monitoring and assessing a Japanese FLES program: Ambiance and achievement. *Language Learning*, 46(3), 497–528.
- Dutchner, N. (1995). *The use of first and second languages in education: A review of international experience*. Pacific Island Discussion Paper Series No.1. Retrieved from <http://documents1.worldbank.org/curated/en/131161468770987263/pdf/multi-page.pdf>
- Education* (pp. 3–15). Multilingual Matters.
- Harrison, W., Prator, C., & Tucker, G. R. (1975). *English-language policy survey of Jordan*. Center for Applied Linguistics.

- In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual*
- Kaplan, R. B., & Baldauf, R. B., Jr. (1997). *Language planning: From practice to theory*. Multilingual Matters.
- Kaplan, R. B., & Baldauf, R. B., Jr. (2003). *Language and language-in-education planning in the Pacific Basin*. Kluwer.
- Ke, S., & Tucker, G. R. (2016). Unleavened cakes: L2 Chinese education for South Asian students in Hong Kong. *K-12 Chinese Language Teaching*. Retrieved from <https://clta-us.org/wp-content/uploads/2015/09/Unleavened-Cakes-L2-Chinese-Education-for-South-Asian-Students-in-Hong-Kong.pdf>
- Lambert, R. D., & Tucker, G. R. (1981, April). *The role of NIE in stimulating innovative language learning and teaching*. Paper presented at a meeting of the National Council on Education Research, Washington, DC.
- Lambert, W. E., & Tucker, G. R. (1972). *Bilingual education of children: The St. Lambert experiment*. Newbury House Publishers.
- Li, S., & Tucker, G. R. (2013). A survey of the US Confucius Institutes: Opportunities and challenges in promoting Chinese language and culture education. *Journal of the Chinese Language Teachers Association*, 48(1), 29–54.
- McField, G. (Ed.). (2014). *The miseducation of English learners: A tale of three states and lessons to be learned*. Information Age Publishing.
- Nair, I., Norman, M., Tucker, G. R., & Burkert, A. (2012). The challenge of global literacy: An ideal opportunity for liberal professional education. *Liberal Education*, 98(1), 56–61.
- Padilla, A. M., Lindholm, K. J., Chen, A., Durán, R., Hakuta, K., Lambert, W., & Tucker, G. R. (1991). The English-only movement: Myths, reality, and implications for psychology. *American Psychologist*, 46(2), 120–130.
- Shohamy, E. (2006). *Language policy: Hidden agendas and new approaches*. Routledge.
- TESOL International Association (TESOL). (2019). *Standards for Initial TESOL Pre-K-12 Teacher Preparation Programs*. Author.
- Tucker, G. R. (1981). Bilingual education research in developing countries: Implications for the United States. *California Journal of Teacher Education*, 8(3), 1–14.
- Tucker, G. R. (1984). Toward the development of a language-competent American society. *International Journal of the Sociology of Language*, 45, 153–160.
- Tucker, G. R. (1986). Implications of Canadian research for promoting a language competent American society. In J. A. Fishman (Ed.), *The Fergusonian impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday Vol 2* (pp. 361–369). Mouton de Gruyter.
- Tucker, G. R. (1987). Educational language policy in the Philippines: A case study. In P. H. Lowenberg (Ed.), *Georgetown University round table on languages and linguistics 1987* (pp. 331–341). Georgetown University Press.
- Tucker, G. R. (1991). Developing a language-competent American society: The role of language planning. In A. G. Reynolds (Ed.), *Bilingualism, multiculturalism, and second language learning: The McGill conference in honour of Wallace E. Lambert* (pp. 65–79). Psychology Press
- Tucker, G. R. (1997). Developing a language-competent American society: Implications of the English-only movement. In T. Bongaerts, & K. de Bot (Eds.), *Perspectives on foreign language policy: Studies in honor of Theo van Els* (pp. 89–98). John Benjamins.
- Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education
- Tucker, G. R. (2001a). A global perspective on bilingualism and bilingual education. In J. E. Alatis & A.-H. Tan (Eds.), *Georgetown University round table on languages and linguistics 1999* (pp. 332–340). Georgetown University Press.
- Tucker, G. R. (2001b). Age of beginning instruction. *TESOL Quarterly*, 35(4), 597–598.
- Tucker, G. R. (2004). Learning other languages: The case for promoting bilinguality within an international community. In D. Resnick, & D. Scott (Eds.), *The innovative university*. Carnegie Mellon University Press.

- Tucker, G. R. (2005). Innovative language education programmes for heritage language students: The special case of Puerto Ricans? *International Journal of Bilingual Education and Bilingualism*, 8(2-3), 188–195.
- Tucker, G. R. (2008). Learning other languages: The case for promoting bilingualism in our educational system. In D. M. Brinton, O., Kagan, & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 39–52). Routledge.
- Tucker, G. R., & Donato, R. (2003). Implementing a district-wide foreign language program: A case study of acquisition planning and curricular innovation. In D. Tannen & J. E. Alatis (Eds.), *Georgetown University round table on languages and linguistics 2001* (pp. 178–195). Georgetown University Press.
- Tucker, G. R., Donato, R., & Antonek, J. L. (1996). Documenting growth in a Japanese FLES program. *Foreign Language Annals*, 29(4), 539–550.
- WIDA. (2014). *2012 amplification of the English Language development standards, Kindergarten–Grade 12*. WIDA.

EFL Literacy Development in Ethnic and Language Minority Learners: Implications from Tertiary-Level EFL Teaching and Learning in Ethnic Minorities in China



Sihui (Echo) Ke

Abstract Guided by the conceptual model of biliteracy development (Dixon & Wu Language Teaching, 47(4), 414–449, 2014), this research aimed to examine tertiary-level English-as-a-foreign-language (EFL) literacy development in ethnic minorities (EM) in China, an underexamined population in the literature. Three research questions (RQs) were posed: RQ1: (a) What challenges do EM students encounter in university EFL learning? (b) How do they cope with the challenges? RQ2: (a) What is the status quo of EFL teaching practices for EM university students in China? (b) Are there any effective EFL models/programs/pedagogies for EM students in China? RQ3: What are EM students' EFL literacy profiles? To answer these questions, two studies were carried out: a synthesis of nine primary studies published between May 2010 and May 2021, as well as an exploratory empirical study comparing English listening and reading comprehension between EM and Han majority students in a major university in Guangdong, China. There were four main findings: first, there were four major challenges in tertiary-level EFL teaching and learning for EM students: an exam-oriented English curriculum, limited English educational resources for EM students, language dilemma (i.e., Chinese is used as the medium of instruction in English classes), as well as affective, sociocultural, and psychological issues. In response to the challenges, EM students adopted different coping strategies, ranging from drawing on their multilingual repertoire as transferable resources for English learning, to ethnic identity revitalization via cultural practice, to more investment in English learning due to external facilitation. Second, there is no effective top-down guidance for tertiary-level EFL teaching practices for EM students, yet bottom-up efforts yielded improvement in EM students' EFL writing. Third, against the stereotype of EM students as “poor English learners”, EM students' EFL listening comprehension was not significantly different from that in

S. (Echo) Ke (✉)
University of Kentucky, Lexington, KY, USA
e-mail: sihui.ke@uky.edu

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_14

301

Han majority students. Finally, there is significant intragroup variability in EM students' English literacy profiles: English reading comprehension is affected by the usage of the EM language writing system; EM students who had functional writing systems of limited usage in their EM languages performed as well as those who had functional writing systems of broad usage, yet weaker than the Han majority students. This chapter provides new data as well as theoretical and methodological implications for future boundary-crossing and interdisciplinary research of language and literacy education in multilingual contexts.

Keywords Multilingual/L3 literacy · English · Ethnic and language minority learner · China

1 Introduction

China has a total population of 1.4 billion, with 12.55 million (approximately 8.89%) belonging to ethnic minority (EM) groups (The State Council for the Seventh National Population Census, 2021). Many of these EM groups have their own languages; not many of these languages, though, have a writing system of broad usage. In the past decades, China has been undertaking one of the largest multilingual engineering projects in the world (Zhou & Ross, 2004). However, previous empirical research about the EM population in China has mainly focused on K-12 level bilingual acquisition (EM language and Putonghua¹), and it is not until the recent decade that more attention has been paid to *Sanyu Jiantong* (mastery of three languages including the EM language, Putonghua, and English) education at the tertiary level (Feng & Adamson, 2015; Wang, 2016).

This research had two aims. The first was to evaluate the status quo of English-as-a-foreign-language (EFL henceforth) learning and teaching for EM university students in China via a synthesis of prior empirical research published between January 2010 and May 2021. Specifically, I focused on the challenges EM students encounter in the tertiary-level EFL education setting and explored effective EFL models/programs/pedagogies for this student group. The second aim was to explore and assess EM students' EFL literacy profiles via an empirical study with students in a major university in Guangdong, China. In the existing literature, EM students are often reported to be socially positioned or even self-positioned as “poor English learners;” yet no empirical studies, to the best of my knowledge, have assessed their English literacy profiles. The findings of this research expand current micro-, meso-, and macro-level understandings of tertiary-level EFL teaching and learning for EM

¹Putonghua was defined at the standardization conference of October, 1955, as the common language of the People's Republic of China; it has the following features: the pronunciation of the general Peking Dialect, the grammar of the Northern Chinese Dialects, and the vocabulary of modern colloquial Chinese literature (as cited in Wrenn, 1975). In this manuscript, “Putonghua”, “Mandarin” and “Chinese” are used interchangeably unless clarification is needed.

students in China and inform EFL-related policy and planning in other multilingual settings.

As indicated by Zhang and Miller in the Introduction of this volume, boundary crossing can occur at many levels, ranging from disciplines to theoretical perspectives, paradigms, methodologies and methods, contexts, languages, programs or learners. It is expected that the findings of this chapter will provide broader implications for future multilingual educational research by combining different methods (i.e., conducting a synthesis study of pertinent quantitative, qualitative and mixed-methods research and an empirical quantitative study); examining macro-, meso- and micro-level contextual effects; and analyzing the literacy profiles of language minority learners of English who use typologically different languages at home.

2 Literature Review

2.1 *Theoretical Framework: The Conceptual Model of Biliteracy Development (Dixon & Wu, 2014)*

In Dixon and Wu's (2014) model, they took an inclusive view of *literacy*, which encompasses not only reading and writing skills but also oral language. Their intention was mapping the micro-, meso-, and macrosystems influencing home language and literacy practices among immigrant second language (L2) learners in different countries. It was hypothesized that the microsystems of the home and school have a *direct* impact on literacy development, as does the mesosystem created by the interaction of home and school; the macrosystem, including societal attitudes and policies toward immigrants, *indirectly* influences a learner's literacy development and how her home resources and practices are regarded, as well as how she might interact with societal practices to impact L2 literacy development. It was also hypothesized that that the relationship between first language (L1) and L2 literacy skills is bidirectional: literacy skills developed in one language help to promote similar skills in the other. Dixon and Wu's model, echoing the theme of this volume, has an inherent emphasis on boundary crossing in understanding biliteracy/multiliteracy development in linguistically diverse contexts. Since Dixon and Wu (2014) focused on children, an emphasis on the home context was more pertinent in their research. In the next section, I focus on the school context for university adult EM learners.

2.2 Context: Tertiary-level English Education for Ethnic Minority Students in China

In a review of English teaching and learning in EM regions in China, Feng (2012) explained the sociopolitical force behind the shift from bilingualism to trilingualism for EM students (i.e., globalization in the wider context and the state-level policy to promote English language education at all levels of educational institutions throughout the country; Chinese Ministry of Education, 2001a, b, c) and outlined four challenges. The challenges include (a) a lack of resources (e.g., a shortage of qualified English language teachers; diverse English provisions in remote/rural EM regions as compared to geographically and economically advantaged areas occupied by the Han majority), (b) the paradox of preferential policies toward EM students in tertiary institution admission, (c) the *zhongjiayu* (medium of instruction) dilemma (English is taught by Chinese-speaking teachers and textbooks are compiled by Chinese intellectuals who use Chinese as the medium of instruction), and (d) socio-cultural or affective issues (e.g., a lack of motivation to learn English, and self-perception as inferior to Han majority counterparts).

Notably, Feng cited Cummins's (1986, 2000) *power relationship* construct and explained that current English language teaching practices using Chinese as the medium of instruction are not theoretically supported because most EM students have not reached a threshold level in L2 Chinese. In spite of the aforementioned challenges he outlined, Feng held that *sanjujiantong*, or trilingualism and trilingual education for EM students in China, may not be as gloomy as many perceived.² He thought the future is encouraging, citing the successful multilingual programs in the Yanbian Korean Autonomous Prefecture in Jilin Province and the recent rapid increase of certified English language educators. One limitation of Feng's delineations is, though, that they focused on EM groups with functional writing systems in broad usage in their EM languages (e.g., Korean, Mongolian, Tibetan, Uygur, Yi). Those groups have received more attention than other EM groups who have limited usage of EM language writing systems or do not have written scripts in their EM languages at all (see Wang & Postiglione, 2008 for a study about the Dongxiang group in the latter case).

²Feng (2012) reviewed Bastid-Bruguier's (2001) *marginalization* perspective. Bastid-Bruguier claimed that the national drive for English language education in China is further empowering the already powerful majority Han group, leaving minority people even further behind.

2.3 Long-lasting Impact of EM Language Writing System Usage on EFL Literacy Acquisition

The impact of EM language writing system usage on EFL literacy acquisition is of both theoretical and pedagogical interest. Theoretically speaking, while there has been a consensus that L1-induced transfer can facilitate second language/L2 literacy acquisition (Chung et al., 2019; Koda, 2005), research evidence about how L1 experience influences third language (L3) literacy acquisition has just emerged recently. Existing literature suggests that L1 metalinguistic awareness affects both L2 and L3 reading subskill acquisition (e.g., L1 Korean-L2 English-L3 Chinese in Cho & Tong, 2014); there is no direct effect of L2 reading on L3 reading, yet it is possible for L2 reading to predict L3 reading indirectly via L1 reading subskills (e.g., L1 Urdu-L2 Arabic-L3 English in Rauch et al., 2012).

Pedagogically speaking, this is highly relevant to the choice of medium of instruction in English teaching for EM students in China. Research evidence seems to support using the EM language as the medium of instruction (as opposed to Chinese) in English teaching. One may argue that the L3 literacy studies reviewed above are not based on data generated from EM students in China. To the best of my knowledge, there is only indirect empirical evidence pertinent to the relative contributions of L1 and L2 experiences to L3 English acquisition in EM student in China (e.g., Cai & Yang, 2010; Li et al., 2008). While it is generally agreed that L1 EM experience has long-lasting impacts on L3 English learning, there have been mixed findings about the effect of L2 Chinese experience on L3 English literacy acquisition. For example, Li et al. (2008) observed that only L3 English grammar knowledge was significantly correlated with L3 English reading in Uygur university students in Xinjiang, and L2 Chinese grammar knowledge did not correlate with L3 English reading. In contrast, Cai and Yang (2010) compared the English reading and writing composite scores on a college entrance exam between two groups of Uygur and Kazakh high school seniors in Xinjiang – an EM-language-Chinese-English trilliterate group and a Chinese-English biliterate group, and found that the biliterate group's performance was stronger. Notably, Cai and Yang's findings were based on reported scores on a college entrance exam; no language assessment was administered by the researchers.

Taken together, emerging evidence has suggested that EFL literacy (even at a minimum level) is a strong empowering tool for EM students' educational and socioeconomic advancement (Wei et al., 2021). However, current knowledge has been mainly generated based on scholars' opinions (e.g., Jacob & Park, 2013) or word of mouth. There is little empirical evidence assessing university EFL learning and teaching in EM students, including these students' English literacy profiles. The gap in research on multilingual (EM language, Chinese and English) literacy development in EM students in China is wide. This gap, however, limits language education practice and policy for EM students.

3 The Current Research: An Overview

This research aimed to evaluate existing empirical evidence of EFL teaching and learning for EM students in China in the past decade and to explore EM students' EFL literacy profiles. Three research questions (RQs) were posed:

RQ1. (a) What challenges do EM students encounter in university EFL learning? (b) How do they cope with the challenges?

RQ2. (a) What is the status quo of EFL teaching practices for EM university students in China? (b) Are there any effective³ EFL models/programs/pedagogies for EM students in China?

RQ3. What are EM students' EFL literacy profiles?

To answer RQs 1 and 2, Study One was carried out by searching for, screening and synthesizing relevant primary research published between January 2010 and May 2021. To answer RQ3, Study Two was conducted to assess the EFL literacy profiles in 547 EM and Han majority students in a major university in Guangdong, China.

4 Study One: Research Synthesis

4.1 Literature Search, Coding, and Characteristics of the Primary Studies

A literature search of two databases (i.e., ProQuest and Web of Science) was conducted using a combination of three sets of key words (i.e., English/foreign language/third language/trilingual/multilingual, ethnic minority/ethnic minorities/minority, and China). The initial searches resulted in 873 hits, which were further screened by a-priori inclusion and exclusion criteria: Studies included were relevant to the synthesis scope, published as peer-reviewed journal articles or book chapters, written in English, and data-driven; studies excluded were irrelevant studies, studies focusing on EM and Putonghua only, as well as duplicated reports. After the screening, a total of nine empirical studies (listed in [Appendix A](#)) were selected, analyzed, and then coded by the author.

Because of the small number of primary studies, qualitative inductive coding was adopted (see detailed coding in the online supplementary material: osf.io/b49ed). Coding categories included author and publication year, ethnicity, research site, research methods (e.g., quantitative, qualitative, mixed-methods), research design (e.g., longitudinal, cross-sectional, interventional), participants, theoretical framework (if any), primary study purpose(s), and findings pertinent to RQs 1 and 2 posed in the present study.

³Effectiveness is evidenced as positive change(s) in EM students' affective, cognitive, linguistic, or sociocultural development due to certain pedagogical intervention.

The nine primary studies covered five EM groups, including Mongolian, Naxi, Tibetan, Uyghur, and Yi, all having written scripts in their EM mother tongues. Most studies ($k = 6$) were conducted in the EM regions whereas three studies focused on EM students in economically more advanced areas. The research method used was predominantly qualitative case study ($k = 8$), with only one quantitative questionnaire survey study (Wei et al., 2021). As to the theoretical frameworks, partially due to the qualitative nature of the majority, most researchers cited such constructs as cultural capital (Bourdieu, 1977), identity (Norton, 2013; Peirce, 1995) and investment (Norton, 1997). Zhang (2018) was the only study that cited linguistic theories (i.e., systemic functional linguistics), and, not surprisingly, it is also the only study that measured an EM student's English literacy ability (writing).

4.2 Results

In response to RQ1a) *What challenges do EM university students encounter in EFL learning?* and RQ1b) *How do they cope with the challenges?*, I was able to identify pertinent evidence in nine and four primary studies respectively. Four major challenges were coded: (1) exam-oriented English curriculum ($k = 1$); (2) limited English educational resources for EM students ($k = 4$); (3) language dilemma (i.e., Chinese is used as the medium of instruction in English classes) ($k = 3$); (4) affective, sociocultural, and psychological issues ($k = 4$).⁴ As to EM students' coping strategies, one study identified that students actively drew on their multilingual repertoire as transferrable resources for English learning (Guo & Gu, 2016); another study observed ethnic identity revitalization via cultural practice (Wang et al., 2020); another two studies found more investment in English learning post researcher-led pedagogical intervention for EM students (Jiang et al., 2020; Zhang, 2018).

Regarding RQ2a) *What is the status quo of EFL teaching practices for EM university students in China?*, this synthesis did not identify any change since the publication of Feng's (2012) review: university EFL teaching is still Han-majority, exam-oriented, and not responsive to the special needs of EM students. No top-down English teaching guidance tailored for EM students has emerged in the process of synthesizing the nine primary studies. Finally, RQ2b) is concerned with effective EFL models/programs/pedagogies for EM university students in China. Two EFL pedagogies (digital multimodal composing in Jiang et al., 2020; and systemic functional linguistics-based teaching in Zhang, 2018) were identified. Details of those pedagogies are presented in the next section.

⁴Since some of the nine studies covered more than one category of "challenge", the totality was bigger than nine.

4.3 Discussion

To recapitulate, the purpose of Study One was to evaluate the status quo of university EFL education for EM students in China. As a comparison to Feng's (2012) review, this synthesis identified both similar and different challenges. First, similar to Feng's finding, language dilemma (i.e., using Chinese as the medium of English instruction in classrooms mixed with EM and Han majority students) is still an outstanding issue. Second, this synthesis also identified affective, sociocultural, and psychological issues in EM students in English classes or their university experiences in general. I grouped these issues in one category because they are interconnected. When an EM student moves from remote/rural mountain areas to a new environment such as a university setting (usually located in the capital city in an EM region, such as Xinjiang, or in other more economically advanced regions, such as Guangdong in the case reported in Study Two below), s/he may have identity issues due to the lack of EM groups on campus (e.g., Ge et al., 2012). S/he could be subject to teachers' and Han majority classmates' stereotyping as "poor English language learners" and even self-position as being inferior to the Han majority (Wang et al., 2020). As a result, EM students' class participation is marginalized, and, even worse, the stereotype of EM students as "poor English language learners" is further perpetuated (Wang et al., 2020). Moreover, university administrators do not pay enough attention to students' (not just limited to EM students') psychological health (Ge et al., 2012).

Two new issues were identified in this synthesis, namely, the exam-oriented English curriculum and limited English educational resources for EM students. Notably, these two issues are intertwining. It is common for universities in China to require undergraduate students (regardless of their ethnic background) to pass a standardized national English test, such as CET4 or a university internal English test, to fulfill graduation requirements. Passing these high-stakes tests is often also considered a prerequisite for students to be enrolled in English enrichment/elective courses or to transfer to a different major. Because many EM students could not meet the English testing requirement, they do not have access to the English enrichment/elective courses and need to retake basic English courses irrespective to EM students' needs. It should be noted that it has been reported in the primary studies that universities inflate scores as part of a preferential policy for EM students (Han et al., 2016). For instance, Han et al. (2016) observed that a university-internal English language preferential policy led to mixed opinions among EM and Han majority students. While most students endorsed the English language preferential policy, some Han majority students from economically less developed regions thought it was unfair, and some EM students were also against it because the preferential policy was perceived to have negative effects on their English learning, possibly lowering their motivation.

Despite the challenges and issues reviewed above, some EM students were found to develop resilience and coping strategies, some being self-regulated (Guo & Gu, 2016; Wang et al., 2020), others with external facilitation (e.g., Jiang et al., 2020;

Zhang, 2018). For example, Guo and Gu (2016) interviewed a participant learning Korean as a foreign language, in addition to the three languages she was juggling. The participant reported that the success in English learning bolstered her confidence and impetus to learn other foreign languages, and she was able to apply cognate awareness across languages (e.g., recognizing Chinese and English loan words in Korean). In addition, some Uygur participants in Guo and Gu (2016) reported that their religious knowledge in their EM language facilitated English learning because of content similarity between the Bible and the Quran. The benefits of cultural practices in developing EM students' linguistic resilience should be acknowledged in this regard. Another example can be found in Wang et al.'s (2020) study. They observed that a Tibetan EM student crushed Tibetan incense and put the powder into a bag so that he could carry the smell because his Han majority dormmates could not stand the smell and did not understand Tibetan culture. The Tibetan EM student had to negotiate sociocultural practices for the sake of Tibetan ethnic identity revitalization. Wang et al. (2020) did not, however, specify how the students' negotiation was related with his EFL learning on campus.

Another important finding of this synthesis is perhaps the insights that emerged from the two interventional studies with bottom-up classroom-based efforts aimed at improving English language teaching practices for individual EM students (i.e., Jiang et al., 2020; Zhang, 2018). In what follows, I provide a brief description of the setting and the EM student participant in each case, describe the regular routines and/or principles of the EFL pedagogy, and evaluate the effectiveness of the pedagogy by comparing pre- and post-intervention performances in the EM student participants.

Jiang et al.'s (2020) interventional study was conducted in a university in a coastal city in China, which recruited EM students through a policy that aims to bridge the educational inequity between Eastern and Western China. The intervention involved an English instructor's implementation of six digital multimodal composing (DMC) projects over two academic semesters for 32 weeks, as part of a larger research project aimed at exploring how technology can be used in teaching. According to Jiang et al. (2020, p. 963), "in a typical DMC project, students generally went through six phases of learning: textbook reading, brainstorming, resource collection and production, video editing, in-class sharing, and disseminating videos through the internet." The focal student participant was a female Tibetan freshman, who grew up and completed primary and secondary education in the Tibet Autonomous Region (TAR). Before receiving the DMC intervention, the student had a phobia toward English classes and was not engaged in class participation. After the intervention, she reported that she enjoyed the experience of group writing, continued independent learning via recording, and was able to speak English in class by using her ethnic knowledge as a resource. Jiang et al. (2020) suggested that the DMC approach was effective and empowering for three reasons: first, DMC gives voice to EM students and provides space for them to challenge the traditional ideology that devalues/ignores ethnic cultures; second, DMC provides the technology affordance that allows EM students to control their pronunciation and overcome English-related phobia; last, the collaborative learning, peer support, and

teacher recognition of cultural differences in the DMC projects also empower EM students.

Notably, Jiang et al.'s (2020) study reviewed above did not assess the EM student's English composing/writing ability. Zhang's (2018) study also involved English writing, but in contrast to Jiang et al. (2020), included linguistic assessment. Specifically, Zhang (2018) examined how systemic functional linguistics (SFL)-based teaching affected EM students' adjustment to academic English writing discourse. It was a case study conducted with a Tibetan female freshman majoring in English in a university in Beijing, China. The student completed primary school education in the TAR, then attended secondary education at a boarding school outside the TAR. She was the only EM student in an English writing class. SFL-based instruction featured teacher-student deconstruction of learning materials in which the teacher guided students in understanding SFL-based constructs (e.g., genre, register, ideational meaning, interpersonal meaning, and textual meaning) and their relation to written discourse and independent learning. Zhang (2018) compared the pre- and post-intervention development by recording the student's academic knowledge development (e.g., cross-linguistic understanding and the understanding of the SFL-based constructs), documenting her use of SFL-based knowledge in providing peer feedback, and comparing her initial draft and final writing assignment. The effectiveness of SFL-based instruction lied in that the EM student not only increased her academic English discourse knowledge, but also became more engaged in interaction with peers and the teacher in class and more confident in providing peer feedback and participating in class discussion.

In summary, with respect to the status quo of tertiary-level EFL teaching and learning for EM students in China, to date, major challenges have remained unaddressed. The fundamental issue is the lack of sensitivity to the needs of EM students at the macro- and meso-level (see Dixon & Wu, 2014). Current preferential college admission policies or university-internal English test score inflation policies did not turn out to be an effective top-down approach. As suggested by emerging research evidence, it was individual student-initiated and teacher-led linguistically and culturally responsive practices at the micro level such as digital multimodal composing or SFL-based instruction, which acknowledged EM students' multilingual repertoire as assets, that fostered EM students' resilience against the stereotype of "poor English learners" and consequently promoted their engagement in EFL learning as well as EFL literacy development.

5 Study Two: Exploring EM Students' English Literacy Profiles

5.1 Method

5.1.1 Setting

Data were collected at a satellite campus of a nationally highly ranked university in Guangdong, China. The university was STEM-oriented, and English was usually learned as a supplementary (non-major) subject. Undergraduate students were enrolled in basic English courses after taking a placement test in their freshman year. They could also select additional enrichment English-medium courses after they fulfilled the basic English course requirement.

5.1.2 Participants

A total of 547 freshmen participated in the placement test. The female to male ratio was 1:3.27. Their age ranged between 18 and 19 years old. They were further categorized into three groups for the purpose of the present study: twenty-seven EM students with functional writing systems of broad usage (including Hui, Miao, Mongolian, Uygur, Yi, and Zhuang); 22 EM students with functional writing systems of limited usage (including Chuanqing, Buyi, Lisu, Manchu, Shui, Tong, Tujia, and Yao); and 498 Han majority students.

5.1.3 Instrument

A university-internal English placement test was administered. It included two 50-point sections: listening comprehension and reading comprehension. There were three types of multiple-choice questions in the listening comprehension section: short news, long conversations, and essays. There were two types of multiple-choice reading comprehension questions: skimming and scanning (one passage with five questions) and intensive reading (two passages with a total of 10 questions). All question items were adopted from retired CET-4 or CET-6 tests.⁵

⁵The College English Test (CET) is a large-scale standardized exam administered by the Ministry of Education in China, including two levels (i.e., CET-4 and CET-6). The CET is widely recognized among Chinese institutions and employers. An English introduction to China's CET is available at: <https://wenr.wes.org/2018/08/an-introduction-to-chinas-college-english-test-cet>

5.1.4 Data Analysis Procedures

In order to answer RQ3, *What are EM students' EFL literacy profiles?*, I conducted a MANOVA with listening comprehension and reading comprehension as the dependent variables, and student group membership (the three groups mentioned earlier) as the independent variable.

5.2 Results

5.2.1 Descriptive Statistics

In general, the average total score (listening and reading comprehension combined) for all three groups was above 60 out of 100 points (as shown in Table 1 and Fig. 1). It seemed from Fig. 1 that there was no notable difference in either listening comprehension or reading comprehension between the EM group with functional writing systems of broad usage and the Han majority group, and that the EM group with functional writing systems of limited usage had lower performance in reading comprehension than the Han majority group.

Table 1 Descriptive statistics for listening comprehension, reading comprehension and composite scores ($N = 547$)

EFL literacy subcomponent	Student group	M	SD	95% Confidence Interval	
				Lower Bound	Upper Bound
Listening comprehension (MSP = 50)	EMLU (N = 27)	32.09	1.51	6.70	35.05
	EMBU (N = 22)	34.04	1.36	7.14	36.71
	CHI (N = 498)	34.66	0.32	7.07	35.28
Reading comprehension (MSP = 50)	EMLU (N = 27)	28.86	1.68	6.70	32.16
	EMBU (N = 22)	33.22	1.52	7.14	36.20
	CHI (N = 498)	34.13	0.35	7.07	34.82
Composite score (MSP = 100)	EMLU (N = 27)	60.96	2.70	12.87	66.26
	EMBU (N = 22)	67.26	2.44	12.19	72.05
	CHI (N = 498)	68.79	0.57	12.67	69.90

Note. MSP maximum score possible, *EMLU* EM students with EM language writing system of limited usage, *EMBU* EM students with EM language writing system of broad usage, *CHI* Han majority students

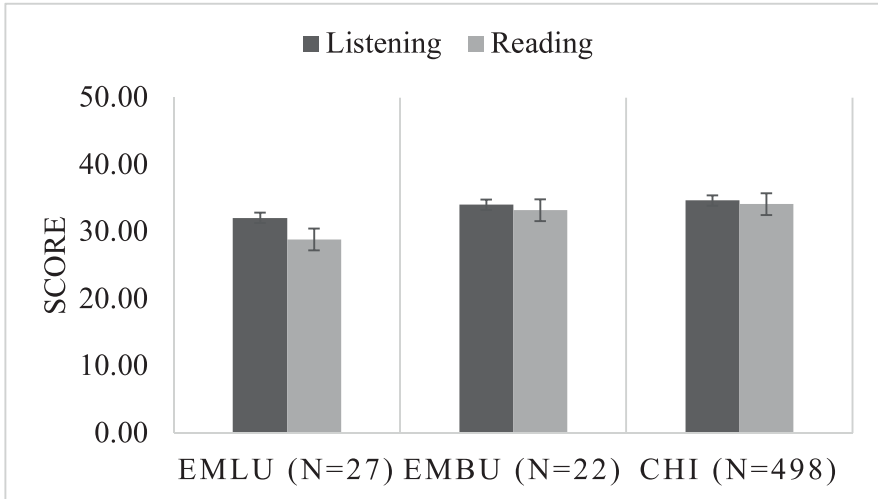


Fig. 1 EFL listening and reading comprehension performances of EM and Han majority students. (Note. EMLU EM students with EM language writing system of limited usage, EMBU EM students with EM language writing system of broad usage; CHI Han majority students)

5.2.2 The Literacy Profiles of EM Students

For MANOVA analyses, I first checked Box’s test of equality of covariance matrices ($F = 0.137, p = .991$) and Levene’s test of equality of error variances (listening comprehension based on mean: Levene statistic = .35, $p = .708$; reading comprehension based on mean: Levene statistic = .49, $p = .611$), and both were nonsignificant. Using Roy’s largest root, there was a significant effect of *group* on students’ EFL literacy profile $\Theta = 0.018, F(2, 543) = 4.89, p = .008$. Furthermore, the results of tests of between-subject effects suggested nonsignificant differences in listening comprehension, $F(2, 544) = 1.47, p = .232$; yet a significant difference in reading comprehension surfaced, $F(2, 544) = 4.81, p = .009$.

Post-hoc pairwise comparison results in reading comprehension are shown in Table 2. The results suggested that there was no significant difference in reading comprehension between any two groups except that the EM student group with EM language writing systems of limited usage had a significantly lower score in reading comprehension than the Han majority group.

5.3 Discussion

There were two major findings from this study exploring the English literacy profiles of EM university students. First, in contrast to the stereotype of “poor English learners” reported in the nine primary studies reviewed earlier, EM students,

Table 2 Post-hoc pairwise comparison results

Dependent Variable	(I) Student group	(J) Student group	Mean difference (I-J)	Std. error	Sig.	95% Confidence interval for difference	
						Lower Bound	Upper Bound
Listening comprehension	EMLU	EMBU	-1.95	2.03	0.338	-5.93	2.04
		CHI	-2.57	1.54	0.095	-5.59	0.45
	EMBU	EMLU	1.95	2.03	0.338	-2.04	5.93
		CHI	-0.63	1.40	0.654	-3.37	2.12
	CHI	EMLU	2.57	1.54	0.095	-0.45	5.59
Reading comprehension	EMLU	EMBU	-4.36	2.26	0.054	-8.80	0.08
		CHI	-5.26*	1.71	0.002	-8.63	-1.90
	EMBU	EMLU	4.36	2.26	0.054	-0.08	8.80
		CHI	-0.90	1.56	0.561	-3.96	2.15
	CHI	EMLU	5.26*	1.71	0.002	1.90	8.63
		EMBU	0.90	1.56	0.561	-2.15	3.96

Note. EMLU students with EM language writing system of limited usage, EMBU students with EM language writing system of broad usage, HI Han majority students

* $p < .05$

especially those with functional EM language writing system of broad usage, performed as well as Han majority students in English listening and reading comprehension. In addition, EM students with functional EM language writing systems of limited usage performed as well as Han majority students in English listening comprehension.

Second, there was significant intragroup variability among EM students due to writing system usage in their EM languages. EM students with functional EM language writing systems of limited usage had lower scores in English reading comprehension as compared to the Han majority students, whereas EM students with functional EM language writing systems of broad usage performed as well as the Han majority students. This finding was consistent with previous third language/multilingual reading literature on the long-lasting impact of L1 literacy experience on L3 reading (e.g., Cho & Tong, 2014; Li et al., 2008; Rauch et al., 2012). In this case, more frequent exposure to EM written language might help EM students develop English literacy skills (e.g., reading comprehension). Notably, however, this finding was different from that of Cai and Yang's (2010) study. To reiterate, Cai and Yang did not find significant influence of EM experience on Uyghur and Kazakh students' English reading and writing skills, which were indexed by their composite scores on the college entrance exam. In contrast, the present study did not just focus on holistic English test scores, but differentiated and measured specific literacy sub-components (i.e., listening and reading comprehension).

Two important implications can be drawn from the two major findings above. First, since EM students and Han majority students seemed to perform equally well in English listening comprehension, EFL teaching in classrooms mixed with EM and Han majority students should be built upon students' oral language foundation so that both student groups can benefit from it. Second, EM students are not a homogeneous group and their previous EM writing system experiences could influence L3 English reading comprehension. Teachers need to take this into consideration when designing differentiated instruction.

6 Conclusions, Limitations, Implications, and Reflections on Boundary-Crossing

EFL literacy is beneficial for EM students' cognitive and psychological development and socioeconomic prospects. Nevertheless, there is little empirical evidence uncovering EFL literacy or multilingual literacy development in EM students in China. In this chapter, I have synthesized research evidence generated in the past decade with respect to EFL learning and teaching for EM university students in China, and identified different EFL literacy profiles within the EM group via a preliminary empirical study. There were four major findings: (1) there were four major challenges in tertiary-level EFL teaching and learning for EM students: an exam-oriented English curriculum, limited English educational resources for EM students, a medium-of-instruction dilemma, as well as affective, sociocultural and psychological issues. In response to the challenges, EM students adopted different coping strategies, ranging from drawing on their multilingual repertoire as transferable resources for English learning, to ethnic identity revitalization via cultural practice, and to more investment in English learning due to external facilitation. (2) There was no effective top-down guidance for tertiary-level EFL teaching practice for EM students; yet bottom-up efforts seemed to yield improvement in EM students' EFL writing. (3) Against the stereotype of EM students as "poor English learners", EM students' EFL listening comprehension was not different from that of Han majority students. And (4), there was significant intragroup variability in EM students' English literacy profiles: only English reading comprehension was affected by the level of usage of an EM language writing system. EM students who have functional writing systems of broad usage in their EM languages performed as well as the Han majority students, and better than those who have functional writing systems of limited usage.

Three major limitations of this research need to be acknowledged: (1) only publications in English have been synthesized. Future synthesis studies should include publications in Chinese as well. (2) The sample in the exploratory empirical study (i.e., EM students in a top-tier university in Guangdong, China) was a convenient one. The findings, therefore, might be not generalizable to EM students studying in universities of lower ranking or located in less developed regions. (3) Only receptive

EFL literacy skills (i.e., listening and reading) were assessed using a standardized test format in Study Two. Future research should include productive skills and researcher-designed assessments as well.

I will now propose an agenda, in light of boundary crossing, for future research on language education policy and planning for EM learners in multilingual settings. First, there is a need to reconsider the proximal theoretical paradigm. In this chapter, I used Dixon and Wu's (2014) model as guidance, which was mainly informed by Bronfenbrenner's bioecological perspective of human development (Bronfenbrenner, 1995; Bronfenbrenner & Morris, 2006). Although the bioecological perspective is advantageous for examining the influence of environment on human development (including literacy development) at macro-, meso-, and micro-level, it has not been applied to researching bilingual or multilingual literacy development. Dixon and Wu (2014) did consider the cross-linguistic influence between L1 and L2 in their conceptual model, yet did not capture the complexities of L3 or multilingual literacy development. Another limitation of the bioecological perspective is that it does not attend to the needs of ethnic and language minority learners, or specify the mechanisms through which different levels of environmental factors jointly influence child development (cf. Coll et al.'s, 1996 integrative model of developmental competencies in minority children).

Second, as Dixon and Wu (2014) pointed out, it is also important for researchers to reflect on their perspectives. Based on the synthesis in this chapter, it seems that there is a dominant deficit perspective of EM students' English literacy development. It is recommended that researchers should consider adopting a strength/resilience-based perspective, not simply comparing EM students against the Han majority students, but rather paying more attention to the intragroup variability, needs, and linguistic resilience (i.e., multilingual literacy versus Chinese-dominant literacy; García-Sánchez, 2019) within the EM group.

Third, methodologically speaking, the primary studies reviewed were dominated by qualitative case studies and cross-sectional designs. There is a need to diversify the research methods and track EM students' development across time. In this chapter, I have combined a synthesis study of qualitative research and a preliminary quantitative study that jointly contributed to the identification of salient personal and contextual factors influencing EM learners' EFL literacy development. It is hoped that this approach will provide implications for the methodological design of future research of language and literacy education policy and planning.

Fourth, as mentioned earlier, the EM population is linguistically and culturally diverse. One linguistic characteristic is the availability and/or the level of usage of an EM language writing system, which has been shown by the evidence of Study Two to have long-lasting impacts on EFL literacy acquisition. Existing research has mainly focused on EM groups with writing systems of broad usage. More attention needs to be paid to EM groups with writing systems of limited usage or even without a written script in their EM languages.

Fifth, systematic design integrating assessment and instruction (versus holistic, standardized, Chinese-medium) is needed to promote sustainable culturally and linguistically responsive practices that recognize EM students' multilingual resources

in EFL teaching and learning. For example, De Angelis (2021) has proposed an integrated approach designed for linguistically and culturally diverse populations that gathers information about the knowledge, skills, and abilities of multilingual learners; uses tools that may be administered in multiple modalities; and uses assessment results that are scored by multilingual examiners and interpreted using data that include information about test takers' language background and living environment.

Finally, I would like to conclude with my reflections on boundary-crossing, which has been defined as “the ability to function competently in multiple contexts” (Walker & Nocon, 2007, p. 178) or “the efforts made to establish continuity [interaction among different communities]” (Bakx et al., 2016, p. 76). Based on the findings of Study One and Study Two, there is obviously discontinuity in existing actions among teachers, researchers, and policy makers. To reach the five future research goals mentioned above, collaboration that crosses boundaries to support EM learners' development of EFL literacy is anticipated among researchers in disciplines within and beyond the field of applied linguistics, teachers, school administrators, and policy makers (Donato & Tucker, 2010; Tucker, 2000).

Appendix A: Nine Primary Studies Selected in the Synthesis

- Ge, W., Tsung, L., & Wing-wah, K. (2012). The pains of becoming trilingual in China: An ethnographic case study of a Naxi college student. *Asia-Pacific Education Researcher*, 21(2), 257–266.
- Guo, X. G., & Gu, M. M. (2016). Identity construction through English language learning in intra-national migration: a study on Uyghur students in China. *Journal of Ethnic and Migration Studies*, 42(14), 2430–2447.
- Han, Y., De Costa, P. I., & Cui, Y. (2016). Examining the English language policy for ethnic minority students in a Chinese university: A language ideology and language regime perspective. *Current Issues in Language Planning*, 17(3–4), 311–331.
- Jiang, L., Yang, M., & Yu, S. (2020). Chinese ethnic minority students' investment in English learning empowered by digital multimodal composing. *TESOL Quarterly*, 54(4), 954–979.
- Liu, J., & Edwards, V. (2017). Trilingual education in China: perspectives from a university programme for minority students. *International Journal of Multilingualism*, 14(1), 38–52.
- Sunuodula, M., & Feng, A. (2011). Learning English as a third language by Uyghur students in Xinjiang: A blessing in disguise?. In A. Feng (Ed.), *English language education across Greater China* (pp. 260–283). Clevedon: Multilingual Matters.
- Wang, H., Chao, X., & Sun, S. (2020). Tibetan students at an interior university in China: Negotiating identity, language, and power. *Asia Pacific Journal of Education*, 40(2), 167–181.

- Wei, R., Jiang, H., & Kong, M. (2021). Attitudes toward trilingualism: A survey study of Chinese Mongolian university students. *Journal of Multilingual and Multicultural Development*, 42(3), 291–306.
- Zhang, X. (2018). Diluting minority students' marginalization in the mainstream college English writing classroom through functional linguistic praxis: A case report from China. *The Asia-Pacific Education Researcher*, 27(6), 465–475.

References

- Bakx, A., Bakker, A., Koopman, M., & Beijaard, D. (2016). Boundary crossing by science teacher researchers in a PhD program. *Teaching and Teacher Education*, 60, 76–87.
- Bastid-Bruguiera, M. (2001). Educational diversity in China. *China Perspectives*, 36, 17–26.
- Bourdieu, P. (1977). *Reproduction in education, society, culture*. Sage.
- Bronfenbrenner, U. (1995). The bioecological model from a life course perspective: Reflections of a participant observer. In P. Moen, G. H. Elder Jr., & K. Lüscher (Eds.), *Examining lives in context: Perspectives on the ecology of human development* (pp. 599–618). American Psychological Association.
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 793–828). John Wiley & Sons.
- Cai, F., & Yang, Z. (2010). Cross-linguistic influence of L2(Chinese) on L3(English) Acquisition of Xinjiang Minority Students [in Chinese]. *Foreign Languages and Their Teaching*, 2, 10–13.
- Chinese Ministry of Education. (2001a). *Guiding ideas to promote English curriculum in primary schools by the Ministry of Education. No. 2*. [in Chinese]. Ministry of Education.
- Chinese Ministry of Education. (2001b). *English curriculum standards for full-time nine-year compulsory education and senior high schools (trial version)* [in Chinese]. Beijing Normal University Press.
- Chinese Ministry of Education. (2001c). *Guiding ideas for improving the quality of undergraduate teaching in higher education institutions* [in Chinese]. Ministry of Education.
- Cho, E., & Tong, F. (2014). Morphology-based cross language transfer from Korean to literacy acquisition in English and Chinese. *Asian Journal of Humanities and Social Sciences*, 2(1), 21–36.
- Chung, S. C., Chen, X., & Geva, E. (2019). Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *Journal of Neurolinguistics*, 50, 149–161.
- Coll, C. G., Crnic, K., Lamberty, G., Wasik, B. H., Jenkins, R., Garcia, H. V., & McAdoo, H. P. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, 67(5), 1891–1914.
- Cummins, J. (1986). Empowering minority students: A framework for intervention. *Harvard Educational Review*, 56(1), 18–36.
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Multilingual Matters.
- De Angelis, G. (2021). *Multilingual testing and assessment*. Multilingual Matters.
- Dixon, L. Q., & Wu, S. (2014). Home language and literacy practices among immigrant second-language learners. *Language Teaching*, 47(4), 414–449.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.

- Feng, A. (2012). English teaching and learning in ethnic minority regions in China: Challenges and opportunities. In J. Ruan & C. B. Leung (Eds.), *Perspectives on teaching and learning English literacy in China* (pp. 129–143). Springer.
- Feng, A., & Adamson, B. (2015). *Trilingualism in education in China: Models and challenges*. Springer.
- García-Sánchez, I. M. (2019). Centering shared linguistic heritage to build language and literacy resilience among immigrant students. In I. M. García-Sánchez & M. F. Orellana (Eds.), *Language and cultural practices in communities and schools: Bridging learning for students from non-dominant groups* (pp. 139–160). Routledge.
- Ge, W., Tsung, L., & Ki, W.-W. (2012). The pains of becoming trilingual in China: An ethnographic case study of a Naxi college student. *Asia-Pacific Education Researcher*, 21(2), 257–266.
- Guo, X. G., & Gu, M. M. (2016). Identity construction through English language learning in intra-national migration: A study on Uyghur students in China. *Journal of Ethnic and Migration Studies*, 42(14), 2430–2447.
- Han, Y., De Costa, P. I., & Cui, Y. (2016). Examining the English language policy for ethnic minority students in a Chinese university: A language ideology and language regime perspective. *Current Issues in Language Planning*, 17(3–4), 311–331.
- Jacob, J. J., & Park, H. (2013). The case of Korean and Mongol minority groups. In G. H. Beckett & G. A. Postiglione (Eds.), *China's assimilationist language policy: The impact on indigenous/minority literacy and social harmony* (pp. 156–172). Routledge.
- Jiang, L., Yang, M., & Yu, S. (2020). Chinese ethnic minority students' investment in English learning empowered by digital multimodal composing. *TESOL Quarterly*, 54(4), 954–979.
- Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.
- Li, L., Mo, L., & Wang, R. (2008). Semantic access of proficient Chinese-English bilinguals to a third language [in Chinese]. *Acta Psychologica Sinica*, 40(5), 523–530.
- Norton, B. (1997). Language, identity, and the ownership of English. *TESOL Quarterly*, 31(3), 409–429.
- Norton, B. (2013). *Identity and language learning: Extending the conversation*. Multilingual Matters.
- Peirce, B. N. (1995). Social identity, investment, and language learning. *TESOL Quarterly*, 29(1), 9–31.
- Rauch, D. P., Naumann, J., & Jude, N. (2012). Metalinguistic awareness mediates effects of full biliteracy on third-language reading proficiency in Turkish–German bilinguals. *International Journal of Bilingualism*, 16(4), 402–418.
- The State Council for the Seventh National Population Census. (2021). *Main data of the seventh national population census*. Retrieved from: http://www.stats.gov.cn/english/PressRelease/202105/t20210510_1817185.html
- Tucker, G. R. (2000). Concluding thoughts: Applied linguistics at the juncture of millennia. *Annual Review of Applied Linguistics*, 20, 241–249.
- Walker, D., & Nocon, H. (2007). Boundary-crossing competence: Theoretical considerations and educational designs. *Mind, Culture, and Activity*, 14(3), 178–195.
- Wang, G. (2016). *Pains and gains of ethnic multilingual learners in China*. Springer.
- Wang, J., & Postiglione, G. A. (2008). China's minorities without written scripts: The case of education access among the Dongxiang. *Journal of Asian Pacific Communication*, 18(2), 166–190.
- Wang, H., Chao, X., & Sun, S. (2020). Tibetan students at an interior university in China: Negotiating identity, language, and power. *Asia Pacific Journal of Education*, 40(2), 167–181.
- Wei, R., Jiang, H., & Kong, M. (2021). Attitudes toward trilingualism: A survey study of Chinese Mongolian university students. *Journal of Multilingual and Multicultural Development*, 42(3), 291–306.
- Wrenn, J. J. (1975). Popularization of Putonghua. *Journal of Chinese Linguistics*, 3(2/3), 221–227.

- Zhang, X. (2018). Diluting minority students' marginalization in the mainstream college English writing classroom through functional linguistic praxis: A case report from China. *The Asia-Pacific Education Researcher*, 27(6), 465–475.
- Zhou, M., & Ross, H. (2004). Introduction: The context of the theory and practice of China's language policy. In M. Zhou & H. Sun (Eds.), *Language policy in the People's republic of China: Theory and practice since 1949* (pp. 1–18). Kluwer Academic Publishers.

“In a Foreign Bubble” in China: Language Use Among International Students During China’s *Belt and Road*



Yi Wang and Wenhao Diao

Abstract This study examines ideologies about language use and policy among international students studying in China against the backdrop of the *Belt and Road* Initiative (BRI). Drawing from ethnographic and interview data collected at an international business program in Shanghai, this chapter investigates how institutional language ideologies become contested in international students’ academic experience in China. Data collection spanned the course of 6 months and included interviews, classroom and participant observations, and artifacts. Findings showed that, due to the belief of English being the international language, students formed a so-called English-speaking “foreign bubble” environment to justify the avoidance of Chinese. Yet, their belief of English as an international language was challenged by the faculty and staff’s assumption that Chinese is necessary for studying in China. These results shed light on the linguistic dilemmas and tensions in the attempt to enhance China’s soft power through study abroad using English.

Keywords *Belt and road* initiative · Study abroad · Language use · English-medium programs · English · Chinese · Language policy

1 Introduction

This chapter focuses on language policy and ideologies in education, an area to which G. Richard Tucker’s scholarship was dedicated. While he was initially trained in psychology, Tucker has frequently engaged in interdisciplinary research to

Y. Wang (✉)

Department of Asian & Asian American Studies, Stony Brook University, Stony Brook, NY, USA

e-mail: yiw.wang@stonybrook.edu

W. Diao

Department of East Asian Studies, University of Arizona, Tucson, AZ, USA

e-mail: wdiao@arizona.edu

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_15

explore social, political, cognitive, and psychological mechanisms and implications of language learning and teaching. From his early work in designing and implementing English language education policy in many parts of the world (e.g., the Philippines) (Tucker, 1968), to his tenure at the Center for Applied Linguistics, and to his more recent endeavors of creating sustainable foreign language programs in K-12 school contexts in the U.S. (e.g., Donato & Tucker, 2007, 2010; Tucker & Donato, 2003), Tucker's research and teaching has dealt with the particularities of bi/multilingual education in specific contexts. Inspired by the legacy of his work that transcends theoretical and methodological boundaries to focus on individual cases and contexts, this chapter draws from an interdisciplinary approach and examines international students in one international business graduate program in China. Specifically, we utilize linguistic and anthropological theories and contextualize the program's policies, ideologies, and practices against the backdrop of China's *Belt and Road Initiative* (henceforth BRI).

Often compared to the Marshall Plan that the U.S. put in place to revitalize the economies of its allies in Europe and elsewhere after World War II, the BRI is a strategy for China to collaborate with other countries through infrastructure building, aid and investment, and trade. The name BRI refers to the land connection between China and Europe (the historical "Silk Road") and the sea trade route linking China to Southeast Asia and extending to Africa (the "Belt") (Gere, 2017). The initiative was first proposed in 2013 by China's President Xi Jinping, who later further advertised it as a "win-win" strategy that could usher in a "new era of globalization" (Xi, 2017). Indeed, the BRI does not simply aim to provide financial and technological assistance to other countries; it is also intended to expand China's economic, cultural, and political influence in the regions. On March 28 of 2015, the BRI's official outlines were released jointly by the National Development and Reform Commission (NDRC), the Ministry of Foreign Affairs (MOFA), and the Ministry of Commerce (MOC) of the People's Republic of China. Since then, over \$900 million has been allocated to this mega project (NDRC, MOC, & MOFA, 2015).

While the BRI is primarily an infrastructure and trade project, education is one crucial component of this initiative. Under the subtitle of "People-to-People Bond" in the BRI outlines, it states, "China provides 10,000 government scholarships to the countries along the *Belt and Road* every year" (Liu, 2018, p. 164). A growing number of students from the BRI countries in these regions are studying abroad in China (National Statistical Reports 2000–2018, MOE). These students often choose to study in China to pursue a career that involves dealing with China. However, an emerging trend among Chinese higher education institutions is to offer courses taught exclusively in English—rather than Chinese—to these international students (Wang, 2020). English-medium academic programs have been a trend growing in many parts of the world (e.g., Coleman, 2006; Rose & McKinley, 2018), but unlike the Englishization of higher education in Europe, where it is mainly intended for domestic students (Coleman, 2006), the programs in China are exclusively designed for international students, and the goal of these programs is mainly derived from the geopolitical and socioeconomic concerns in the context of the BRI (Wang, 2020).

In this study, we focus on one such program that aimed to improve its students' knowledge of China via English. We aim to answer the following two research questions: (1) How do students and other stakeholders of the program perceive languages in use in the English-medium programs in China? (2) What kind of tensions and negotiations, if any, exist between the students and the stakeholders regarding language(s) in use? In what follows, we begin with a brief overview of research on language ideologies in the context of international education. We then introduce the sociopolitical context of the current study and the embedded ideological assumptions in educational policies for international students in China. After a portrait of the site, the participants, and the data generation procedures, the chapter presents findings that illuminate tensions among different social actors in light of the BRI in China and its implications for educators and policymakers. By examining study abroad students in an English-medium program in China negotiating the use of both Chinese and English, our current study breaks traditional lines in the field of applied linguistics that are defined by language, such as ESL or the learning of Chinese as a second language, and we will return to this point of boundary crossing in our Discussion section.

2 Language Ideologies in Study Abroad

Language ideologies are “sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use” (Silverstein, 1979, p. 193). Such rationalization and justification for using a given language are inherently related to political and moral interests within a cultural setting, and therefore language ideologies are fundamentally about the linkage between language and other dimensions of social life (McGroarty, 2002). In other words, language ideologies are the “nexus of language, culture, and politics” (Irvine, 2012, p. 47). The close connection between language ideology and the cultural and political dimensions of social life lends itself to the current inquiry into language use in a particular sociopolitical moment—international education in China in the time of the BRI.

Moreover, language ideology is also particularly relevant to international education because study abroad requires the investment of time and resources, and thus it requires ideological justification (Kubota, 2016). For instance, the global expansion of capitalism over the last several decades has given rise to neoliberal language ideologies (Heller & McElhinny, 2017), and elites who benefit from global capitalism also benefit from multilingualism (e.g., De Costa, 2016). These neoliberal language ideologies have been frequently (re)produced and (re)distributed in the recruitment of study abroad students, and they also organize who should/could study abroad, where they go, and what language(s) they should study while being abroad. Under the framing of neoliberal ideologies, the social imaginary of study abroad promotes a view that positions wealthy Anglophone countries as the center of English use (Kubota, 2016); the desire to learn English has been a dominant force that organizes elite study abroad of this type. For instance, because of the neoliberal

language ideology that equates English with professional opportunities in the global marketplace, many young South Korean children from elite families or families with socioeconomic aspirations study abroad in wealthy Anglophone nations and participate in the learning of English (Park & Bae, 2009; Song, 2010). Growing numbers of Chinese students who attend schools in North America and elsewhere and learn English are also manifestations of a similar ideology (e.g., De Costa, 2016; Gao, 2006).

Meanwhile, language ideologies are multiple and can be contentious with one another (Woolard, 2021), with each functioning as part of a larger and complex web of ideologies that are continuously (re)formed in the process of individuals' lifelong and life-wide experience (Guardado, 2018). Study abroad and language learning is also not only motivated solely by neoliberal ideologies. American study abroad, for instance, promotes neoliberal elite multilingualism (Diao, 2021) that has its roots in the "Grand Tour" ideal of the British gentry class (Gore, 2005). This historically framed idea of study abroad leads to the view that study abroad (and by extension learning other cultures/languages) is a less serious academic endeavor, and it is mostly for wealthy college women to learn Western European languages (Gore, 2005; Kinginger, 2008). Yet simultaneously, study abroad is also sometimes encouraged through a nationalistic lens. For instance, the American promotion of studying abroad in China and the Arab world and the framing of Arabic and Chinese as "critical languages" redistribute orientalist and colonialist ideologies that categorize what constitutes the political and cultural East in relation to the United States (Diao & Trentman, 2016; Trentman & Diao, 2017).

The multiplicity of language ideologies also means that the beliefs and values towards language(s) that study abroad students subscribe to may not align with institutional assumptions or practices (De Costa, 2016; Diao & Trentman, 2016). The discrepancies between students' and institutions' ideologies can manifest themselves as tensions in everyday life and impact the individual experience of language learning and use during students' sojourns overseas (Park & Bae, 2009; Surtees, 2016). Ideologies about language and language learning are socially constituted (Douglas Fir Group, 2016; Kubota, 2016), and therefore investigating tensions in language ideologies in study abroad requires us to situate them in their sociohistorical moment. We now turn to a review of institutionalized language ideologies related to international education that justify and reject English use in China today.

3 Language Ideologies and Study Abroad in China

The adoption of English as the medium of instruction is not uncommon in many countries where English is not a dominant language in society (Kubota, 2013; Robertson & Kedzierski, 2016), such as Western Europe (Coleman, 2006) and Japan (Rose & McKinley, 2018). Yet, the adoption of English in education often triggers the "language ideological debate" (Blommaert, 1999, p. 8), because English

sometimes becomes seen as a potential threat to the status of the national language(s) and, by extension, the cultural identity (Smith, 2004). Although English is a mandatory foreign language subject for domestic students from elementary to postsecondary schools throughout China, to the extent that it has often become synonymous with *waiyu* (the foreign language) (see a similar case in Japan described in Kubota, 2008), recently proposed reforms to reduce the required portion of English in the college admission process reveal the ongoing ideological tensions between English and Chinese in the education system (Gil & Adamson, 2011; Pan & Seargeant, 2012). Moreover, the teaching of English in China is an integral part of the larger ideological structures that aim to produce the “good student” who conforms to China’s institutional goals of nationalism and modernity (Pérez-Milans, 2012). Under such circumstances, Chinese students are required to learn English and simultaneously prohibited from being enrolled in degree programs taught in English only.

Yet English-medium programs still widely exist in China’s higher education—exclusively for international students. The forces of neoliberalism and political concerns are both present in ideologies that promote the use of English in international education in China. The emergence of English-medium programs coincided with China’s increasing participation in economic globalization. When China joined the World Trade Organization in 2001, the state newspaper *People’s Daily* published an article urging the establishment of English programs to increase the number of international students. It argued that by removing the Chinese language, programs can “sweep off the language barrier (扫清留学生的语言障碍)” and “eliminate international students’ concerns about insufficient Chinese language (打消留学生未过语言关的顾虑)” (Dong, 2001). Later in 2010, China’s Ministry of Education released *The Outline of the National Plan for Medium and Long-Term Education Reform and Development (2010–2020)*, which specified the goal of making China the largest study abroad destination in Asia by number of international students by 2020. This series of policies assumes that increasing the number of international students should be the primary goal of China’s international education in a globalized world, and Chinese is an intrinsic obstacle toward promoting study abroad in China to more students overseas.

The BRI, which began in 2013, further made increasing the number of students studying in China a national priority. Previously, China’s approach to enhancing its soft power and culture overseas was through promoting the teaching and learning of the Chinese language (Liu & Tao, 2012), such as establishing the Confucius Institutes (Hubbert, 2019; Li & Tucker, 2013). This approach has become more multilayered in the age of the BRI. In addition to sending teachers and resources overseas, today the Chinese government also encourages international students to study in China. Often referred to as *bringing in* (请进来), this strategy was made explicit in a comment by the Minister of Education, Chen Baosheng, at a press conference in 2018:

Our approach [to international education] has always been simultaneously focusing on “bringing in” and “going global.” ... [Through such collaboration] we can foster talents of various kinds that China and other collaborating countries would both need, and who can

serve as ambassadors of friendship for both sides. This is the strategy that we have to “bring people in.” (Chen, 2018. Originally in Chinese. Translated by the authors.)

Thus, international education becomes a means of “fostering” pro-China professionals; increasing the number of international students in China, it thus directly aligns with the intended goal of the BRI to enhance China’s impact regionally and globally (Wang, 2020). In this context, English is again promoted as an effective and efficient tool to achieve this goal. In another governmental policy document *Liuxue Jihua* (“Plan for Study Abroad Students in China”), the solution to this is to offer “English-taught degree programs” that have “Chinese characteristics.” The term “with Chinese characteristics” was first coined by Deng Xiaoping to introduce a market economy to China’s then Soviet-like system of governance. Since then, it has been frequently recontextualized to justify pragmatic compromises in China’s domestic and foreign policies. In the context of international education here, it rationalizes the use of English as a more efficient medium to promote the Chinese culture and defines a combination of China-centric content and English language as the prominent feature of study abroad in China during the BRI.

Under these policies, China became Asia’s largest study abroad destination by number of students in 2017, with more than 489,000 students studying abroad in China, a 10% rise compared to 2016 (Ministry of Education of the People’s Republic of China, 2018). The number of students studying in China from India, Indonesia, Laos, Pakistan, and Thailand combined—all countries covered by the BRI—has increased more than 20% on average from 2016 to 2017, and students from African and European countries have also increased (Wang, 2020). Prior to 2010, few English-medium programs existed in China; yet in 2018 alone, over 1400 English-medium programs were in operation throughout China, most of which were at the graduate level. This growth of English-medium programs in present-day China thus is derived from both neoliberal ideals that link English to economic globalization and, more importantly, China’s geopolitical and socioeconomic priorities in the era of the BRI.

However, ideologically treating English as a superior medium to promote the Chinese culture and identity to international students can create tensions in practice. Dominant English-speaking countries, such as the U.S. and its closest allies (the U.K., Canada, Australia), have either been actively opposing or hesitant in joining an initiative that would rival the dominance of the U.S. Thus, as aforementioned, the BRI’s focus has largely been countries in Asia, Europe, and Africa, most of which are not Anglophone. Most of the international students coming to China therefore do not speak English as their first language. Furthermore, while English is often promoted as the ideal language to attract more international students, English-taught degree programs are prohibited for Chinese college students. International students in these programs are thus separated from their Chinese peers in both recruitment and daily operation. Separate college admission practices can produce unintended ideological stigmatization related to language and identity; for example, the *kikokushijos* in Kanno’s (2003) study, who were children of Japanese expatriates returning from overseas for college, were perceived to be less qualified

linguistically and academically than their locally raised counterparts. In the case of international students in English programs in China, the separation would continue even after the admission, rendering further questions regarding ideological processes that connect language with identity. The current study was designed to uncover tensions between these policies and their underlying assumptions and students’ language use in everyday life.

4 Methodology

4.1 Site

This six-month ethnographic study was conducted in 2019 at a public university located in the Shanghai, China metropolitan area. The focal site was a two-year, English-medium master’s program in international business which catered exclusively to international students (defined as students who were not citizens of China).

The program was established in 2012 in response to the government’s goal of increasing the number of international students (Fieldnotes, February 25, 2019). It initially sought to recruit students who had been educated in wealthy Anglophone countries or English-degree programs in other developed countries, who could both fund their studies in China and would have high-level academic proficiencies in English (Fieldnotes, March 2, 2019). Of its first cohort that was recruited in 2013, over 90% were self-funded. However, the intended profile of its prospective students quickly shifted after the implementation of the BRI. At the time of the study, a growing number of students came from BRI countries, and many students had received scholarships related to the BRI to study in China (see Participants for more detail).

The curriculum of the program was designed in line with a China-centric focus. It contained a series of featured courses related to “doing business in China,” all of which were taught in English. In addition, a Chinese culture course was also required for the international students enrolled. Although the Chinese language was not required for admission, the students had to take one required Chinese language course if they had not yet passed Level III of the *Hanyu Shuiping Kaoshi*, the elementary level of China’s standardized Chinese proficiency test.¹

¹ Test takers who are able to pass the HSK (Level III) can communicate in Chinese at a basic level in their daily, academic and professional lives (retrieved from <https://www.chinaeducenter.com/en/hsk/hsklevel3.php>). For students who apply for Chinese-medium programs, more than HSK (Level V) is required. Test takers who are able to pass the HSK (Level V) can read Chinese newspapers and magazines, enjoy Chinese films and plays, and give a full-length speech in Chinese.

4.2 Participants

4.2.1 Student Participants

The student participants came from the cohort of students who were admitted in September of 2018. A total of 41 students were enrolled, 30 of whom agreed to participate in this study and completed the interviews. Table 1 presents the background information of the 30 participants, and all names in this study are pseudonyms. Of the 30 participants, 10 were female, and 20 were male. As shown in Table 1, they were from 17 countries, including 17 students who came from

Table 1 Basic information for student participants

No.	Pseudonyms	Age	Gender	Country of origin	Continent
1	<i>Sophia</i> ^a	24	F	Thailand	Asia
2	<i>Tony</i>	25	M	Italy ^b	Europe
3	<i>Venessa</i>	27	F	Singapore ^b	Asia
4	<i>Ploy</i> ^a	25	F	Indonesia ^b	Asia
5	<i>Tiffany</i>	25	F	USA	North America
6	<i>Williams</i> ^a	25	M	Italy ^b	Europe
7	<i>Jacob</i> ^a	24	M	Germany	Europe
8	<i>Theodore</i>	30	M	Germany	Europe
9	<i>Asif</i>	25	M	Kazakhstan ^b	Asia
10	<i>Max</i>	23	M	Czech republic ^b	Europe
11	<i>Farhan</i>	25	M	Indonesia ^b	Asia
12	<i>Zahid</i> ^b	24	M	Niger ^b	Africa
13	<i>Una</i> ^a	25	F	Germany	Europe
14	<i>Carlos</i> ^a	25	M	Brazil	South America
15	<i>Joel</i> ^b	30	M	Brazil	South America
16	<i>Oscar</i>	25	M	Brazil	South America
17	<i>Gerald</i> ^a	24	M	Italy ^b	Europe
18	<i>Larry</i> ^a	24	M	Germany	Europe
19	<i>James</i> ^a	25	M	USA	North America
20	<i>Peter</i> ^a	25	M	Serbia ^b	Europe
21	<i>Ada</i> ^a	25	F	Moldova ^b	Europe
22	<i>Teresa</i>	25	F	France	Europe
23	<i>Chai</i>	29	F	Thailand ^b	Asia
24	<i>Hana</i> ^a	25	F	South Korea ^b	Asia
25	<i>Adrien</i>	25	M	France	Europe
26	<i>Hyungun</i> ^a	26	M	South Korea ^b	Asia
27	<i>Javier</i>	27	M	Serbia ^b	Europe
28	<i>Stefan</i> ^a	23	M	Netherlands	Europe
29	<i>Jiahui</i>	25	F	Canada	North America
30	<i>Alex</i> ^a	22	M	Myanmar ^b	Asia

^aDenotes students who received scholarships

^bDenotes countries in BRI

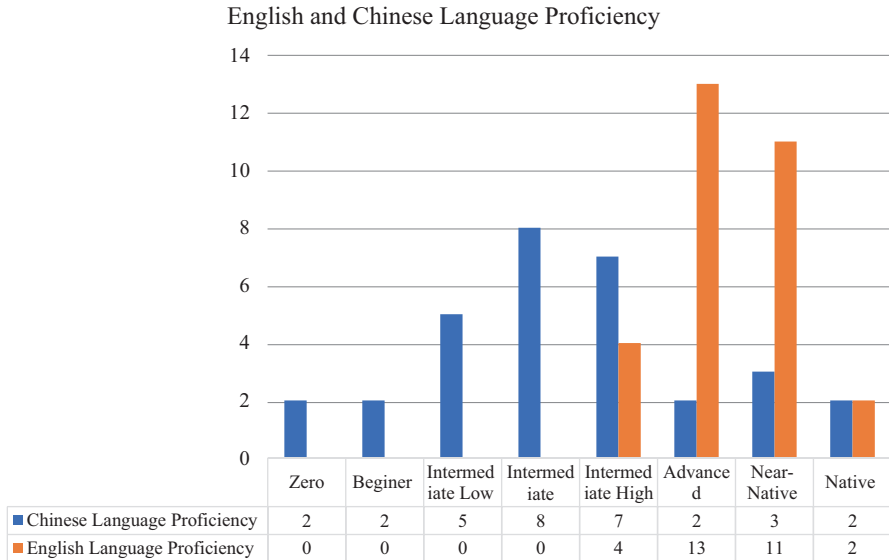


Fig. 1 English and Chinese proficiency level of the student participants

countries that had joined the BRI at the time. While six students were ethnically Chinese (Sophia, Tony, Venessa, Ploy, Williams, and Jiahui), none were Chinese citizens (and were therefore categorized as “international students” in China).

Overall, the students as a group reported higher proficiency in English than Chinese. As shown in Fig. 1, their self-reported English level ranged from intermediate-high to native, with the majority being advanced and near-native. Two of them (James and Tiffany) were originally from the U.S. and were native English speakers. Meanwhile, there was much more variance in the distribution of their Chinese proficiencies. While the majority (20) identified as intermediate Chinese speakers (from intermediate-low to intermediate-high), four participants reported having no (zero) to very little (beginner) Chinese, and five students self-identified as advanced or near-native Chinese speakers. Two participants, Jiahui and Venessa, were born in China and emigrated overseas when they were teenagers, and they considered themselves native speakers of Chinese.

4.2.2 Staff and Faculty Participants

Eight faculty and staff members participated in this study. Table 2 lists the basic information for the stakeholder participants.

All the faculty members who taught courses in English had done academic work in Anglophone countries. For example, Wei was a native Chinese speaker who had received his doctoral degree in the U.S. Takana, who was from Japan and thus the only faculty member who was not a Chinese citizen, obtained his doctoral degree in

Table 2 Profile of stakeholder participants

No.	Pseudonyms	Gender	Citizenship	Position in the program
1	<i>Takana</i>	M	Japan	English course instructor (spring semester)
2	<i>Wei</i>	M	Chinese	English course instructor (spring semester)
3	<i>Yan</i>	F	Chinese	Chinese language instructor (spring semester)
4	<i>Li</i>	F	Chinese	English course instructor (fall semester)
5	<i>Min</i>	F	Chinese	Program director
6	<i>Linfei</i>	F	Chinese	Program administrator
7	<i>Guoqiang</i>	M	Chinese	Program administrator
8	<i>Bai</i>	M	Chinese	Program administrator

London and had lived in China for 3 years. He self-identified as having intermediate-level Chinese language proficiency. One Chinese language instructor, Yan, also participated in this study. She was the instructor of the business Chinese language course, which was optional for the students.

4.3 Data Collection

The first author collected the data in this study from the following sources: (1) a demographic information questionnaire (37 students), (2) monthly semi-structured interviews with student participants (30 students, 75 h of audio recordings), (3) semi-structured interviews with faculty and staff members (8 members, 12 h of audio recordings), (4) field notes and field diary, and (5) artifacts from the students and the program. Participants were given the option to choose the preferred interview language between Chinese and English. Interviews conducted in English were transcribed using the online software *Descript*. Interviews conducted in Chinese was transcribed using the online software *Xunfei Tingjian*. The digitally generated transcripts were then checked and edited for accuracy. The first author also attended and recorded most of the class sessions and social/academic events in the program throughout the semester. Field notes were taken during the observations of classes and events.

4.4 Analysis

All data were imported into NVivo 12, a qualitative analytical software tool. To analyze tensions in language ideologies, a grounded theory (Hadley, 2017) approach was used to categorize recurring themes that emerged from all qualitative data (i.e., interviews, artifacts, and field notes). The grounded theory approach is a bottom-up way to identify emerging ideas, themes, and topics in the data (Hadley, 2019). In line with the grounded theory approach, focused coding (Saldaña, 2016) was

employed as a technique after initial themes emerged. In this process, critical moments in which language ideologies became relevant or even explicit were further coded and analyzed. Discourse and critical discourse analysis techniques (Gee, 2004; Wodak, 2015) were also used to explore further the tensions in language ideologies that emerged.

5 Findings

Three themes of tension emerged in the data, and they are presented as follows.

5.1 *Center of Internationalization: China or West?*

The first theme of tension was related to the interpretation of “international.” This tension between whether being “international” should focus on China or the West is evident in the display of the national flags and the interpretation of it among the students. As shown in Fig. 2, flags representing the nationalities of the students in the college were displayed in the central area of the college building.

These national flags signaled the institute’s goal of internationalization. Each flag represented the national origin of a student; for instance, when a new student (pseudonym Peter) arrived from Serbia, the Serbian flag was added to the display. According to the staff participants, the display of flags was meant to create an international atmosphere as well as inclusiveness and diversity (包容性和多样性) in the



Fig. 2 International flags displayed at the main area of the college

college (Interview with Linfei, April 19, 2019). However, a notable exception was the Chinese flag. Although no student of Chinese citizenship could enter the program, the Chinese flag was positioned in the very center. National flags are a visual illustration of what Billig (1995) refers to as “banal nationalism;” that is, flags function as an everyday reminder of nations that are bounded by geopolitical borders. These national flags, with the Chinese one in the center, visually presented China as the center of the institutional effort of internationalization. Indeed, from the perspective of the program administration, they were located physically in China, and the content was supposedly focused on China.

The presence of the Chinese flag was, however, a mystery to many students, as the Serbian student Peter commented, “The funny thing is we have the Chinese flag here right in the middle, but we do not have Chinese students in our class” (Fieldnotes, March 15, 2019). Students in the program perceived the fact that Chinese students were not included to be a strength and even a defining feature of their program (Fieldnotes, March 27, 2019), and they believed that was the reason behind the use of English in a program that is physically located in China. The students’ interpretation of the flag display was, therefore, that it symbolized the program’s “international” feature, with people “from all over the world” (Fieldnotes, March 27, 2019). To the students, the display of the flags contributed to the “international” imagery of the study abroad experience (Kubota, 2016). The contrast between the Chinese flag in the middle and students’ perception of it being “funny” describes the contradiction between the program’s view and the students’ sentiment.

Moreover, even though the program was intended to teach Chinese business culture, several students described a “Western culture” in the program. Excerpt 1 is an example in which one Nigerian student, Zahid, described “the Western culture” in the program.

Excerpt 1

Zahid: Yeah, the program has the Western culture. You see, the textbooks they are using are all from the U.S., right? Students are from Western countries as well.

Researcher: What about the students from Asia?

Zahid: They are all Westernized, you know? Like, Sophia, Hana, they all studied in a Western country before.

Researcher: Not all of them, right?

Zahid: Yes, but I mean, they chose to study in an English program. (2nd interview with Zahid, April 5, 2019)

At the time of the study, almost half (46.7%) of the student participants were from Europe. The heavy presence of European students was a combined outcome of the program’s initial design to attract affluent students from Western (ideally Anglophone) countries and its later switch to align with China’s BRI. To compromise between the two, the program decided to mostly recruit students from European countries such as Italy and Serbia, who were both assumed to be familiar with academic English because of their European origin and come from the BRI countries (Interview with Guoqiang, March 5, 2019). However, the students interpreted the

preference of recruiting students from Europe and the adoption of English as a sign of Westernization. As pointed out by Zahid here in Excerpt 1, the fact that most students in the program either came from Europe or had studied abroad in these Western countries created a “Western culture” among the students, which did not align with the institutional goal of promoting the Chinese identity and culture by attracting international students to study in China.

5.2 *Taught in English: Academically Superior or Compromised?*

Another circulating ideology among the students was that they were academically better prepared than their Chinese peers who could only study in Chinese programs. Ploy, a student from Indonesia, previously studied in an undergraduate program in China majoring in business that was taught in Chinese. She switched to the English-medium program for her master’s degree while remaining in China. In her first interview, she rationalized the decision by giving two reasons: (1) demographically, students in the English program all would “have international backgrounds” and (2) linguistically, “English is the language for business” (February 25, 2019). In her description, students in the English program were trained in an American academic style and participated in class discussions in a similar way that students in the U.S. would.

Excerpt 2

I think the most significant difference is that the students there [in Chinese programs] are not as open-minded and diverse as we are here. I heard that they [Chinese students] are—just sit there and listen to the lectures, but here you see in our class, we raise questions, do discussions, like, the U.S. style. Also, professors in the English programs all have overseas backgrounds. (...) I mean, we are learning international business and most of, like, the cases, like Harvard cases, are all in English. So, we are more focused on the global not just China, I think. (3rd Interview with Ploy, May 7, 2019)

In Excerpt 2, according to Ploy, students like herself who were in the English program participated differently than those in a Chinese program, and their classroom discussions bore more resemblance to the U.S. educational system—which was perceived to be Anglophone and academically more “global” and therefore superior.

Moreover, as revealed in Zahid’s comment earlier that “the textbooks they are using are all from the U.S.” (Excerpt 1), the students perceived the program’s quality to align with that in the U.S. They frequently cited their textbooks from the U.S., English-speaking faculty, and its resources as features that made their program distinctive from Chinese-taught programs (Fieldnotes, April 5, 2019). By adopting English as its medium of instruction to increase the number of international students, the program had to confront the reality that virtually all textbooks for

business students published within China that focused on China were written in Chinese. The resulting decision was to look overseas and search for an alternative in English, and the eventual materials selected were those published in the U.S. The textbooks from Harvard Business School were selected because they included case studies from China; in actual teaching, the faculty members would include more Chinese cases and focus on multinational corporations' localization processes in China (Fieldnotes, April 20, 2019). These attempts were meant to highlight the program's China focus; however, as revealed in Zahid's quote, the students overlooked the addition of Chinese cases and equated the textbook from the U.S. as evidence of not only Westernization but also a reason to position themselves within a global hierarchy of value and quality that prioritizes the West (particularly the U.S.) By linking English with the U.S., English also became legitimized as the language of power in academic and professional settings (Heller & McElhinny, 2017). In other words, it was not because they were studying in China during the BRI, but rather because they were in an English taught program using American textbooks engaged in U.S.-like classroom interactions, that made them feel they had an advantage in the future job market (Fieldnotes, April 19, 2019).

But the program's staff expressed a sentiment that was remarkably different. From the staff's perspective, English was simply a recruitment strategy to increase the number of international students, because proficiency in academic Chinese was seen as unattainable for most international students interested in business. Thus, similar to Kanno's (2003) observation regarding the stigmatization of Japanese returnees, staff members perceived these students in the English program as linguistically and academically compromised, especially when compared to the Chinese students in the regular programs taught in Chinese (Fieldnotes, April 17, 2019). They believed that these students had to choose the program taught in English because they were linguistically and academically unprepared for a regular program taught in Chinese.

5.3 An English Monolingual “Foreign Bubble”?

The metaphor of “foreign bubble” was frequently used to answer the question “what languages do you use on campus” in interviews. For instance, James, an African American student, responded: “we are exclusively in a foreign bubble” with “no Chinese (students) even allowed” in the program (Interview with James, March 8, 2019). The comment of living in a “foreign bubble” reveals an ideological linkage between language (Chinese versus English) and speaker's citizenship (Chinese versus everyone else). In the interview with Joel, a Brazilian who described his own Chinese language as intermediate, resonated this sentiment and described an exclusively English-speaking “foreign bubble” in which other international students were also perceived to know little to no Chinese:

Excerpt 3

[B]ut also because we have the language barrier, because not all the students can go to a real Chinese seminar, for example, to learn about something we cannot understand. So then I think it's kind of normal for the program to create this bubble, because they can seat you in, like, where you can attend the classes you can understand what they're saying. I think it's kind of normal. (2nd Interview with Joel, April 2, 2019)

Joel's description highlighted the way language organized their social life; (extra) curricular events organized in the Chinese language served the Chinese students, and the events organized in English aimed at international students and excluded Chinese students. The “foreign bubble” thus referred to a space in which non-Chinese citizenship automatically meant the use of English. This “foreign bubble” that contained no Chinese citizen also highlighted the relationship between the linguistic choice and the institutional separation between the international students and the local Chinese students. In a later discussion in the interview, Joel explained that the “foreign bubble” was limited to the university setting, where local Chinese students were not allowed in degree programs taught in English. Stepping out of the university, they would immediately face a “real” China with Chinese citizens who spoke Chinese. Thus, while all Chinese citizens were speakers of Chinese (and often simultaneously presumed to be poor speakers of English), English became the language that connected everyone of non-Chinese citizenship who was a “foreigner” in China.

However, even in this “foreign bubble” on campus that the students perceived to be exclusively English-speaking, the use of the Chinese language was frequently expected. The university administration operated almost entirely in the Chinese language. For instance, the program used an online learning management system, Blackboard, for distributing readings and course materials, and it was completely in Chinese. The students were frequently directed to the website for course-related information, but they were given no additional instructions in English on how to navigate the site. In fact, even the most fervent advocate for English among the program's faculty and staff members expected the students to be able to speak and read some Chinese. For example, Takana, the Japanese professor, made the rule that English was the “official language” in his course (Fieldnotes, March 5, 2019) and stated that all academic communication in the program should take place in English only (Fieldnotes, March 5, 2019). However, during one of his first class meetings, he realized that most of the students could not read any Chinese characters and commented with a surprised tone, “I assume part of the reason you come to China is the Chinese language, so get used to the characters” (Fieldnotes, March 5, 2019). He later also expressed shock when he heard that some students were linguistically unable to navigate Blackboard or access other university resources (Fieldnotes, March 5, 2019).

Professor Takana's comments reveal the institutional assumption that students who came to study in China would inherently have some interest and knowledge regarding the Chinese language and culture, even when they were in an English-medium program that specifically required no Chinese and only recruited students from overseas. This assumption was evident in the Chinese-only online learning systems; it was also manifested repeatedly in day-to-day classroom interactions.

Faculty and staff members frequently inserted Chinese expressions or words without English translations, assuming the students would have no difficulty understanding them. In another course, *Doing Business in China*, the course professor routinely switched from English to Chinese when she discussed Chinese business culture in depth, such as the notion of *mianzi* (“face” in Chinese) in handling business relations (Fieldnotes, March 10, 2019). In other cases, the faculty members in the program would bring up terms in Chinese and ask the students who were Chinese heritage speakers to translate them into English for the rest of the class (Fieldnotes, March 4, 2019).

Many students did not meet such institutional expectations of knowing Chinese. Although the students who had not passed the HSK test were required to enroll in the beginner Chinese language course for one semester (32 class hours), the deliberate separation of the Chinese language from these English-taught courses in the program resulted in students’ lack of motivation to learn Chinese. Many of them could only use the language for simple speech acts such as greetings, apologies, etc. The tension between students’ expectation of an English-only “foreign bubble” and the institutional expectation of their Chinese knowledge became particularly evident in one episode of classroom interaction, when the professor was expected to arrive late. The teaching assistant made the announcement on the board all in Chinese, “The professor will come soon. Please wait.” Figure 3 below was from the scene.

Although the teaching assistant was fluent in English, no English translation was provided in either speaking or writing. Theodore, a student from Germany, commented loudly in English, “Are we supposed to know what this means?” But the comment was ignored by the teaching assistant (Fieldnotes, April 21, 2019). Indeed, even what the students perceived to be an exclusively English-speaking “foreign bubble” was transient. Although their knowledge of Chinese was neither required nor encouraged, it was frequently expected and assumed.

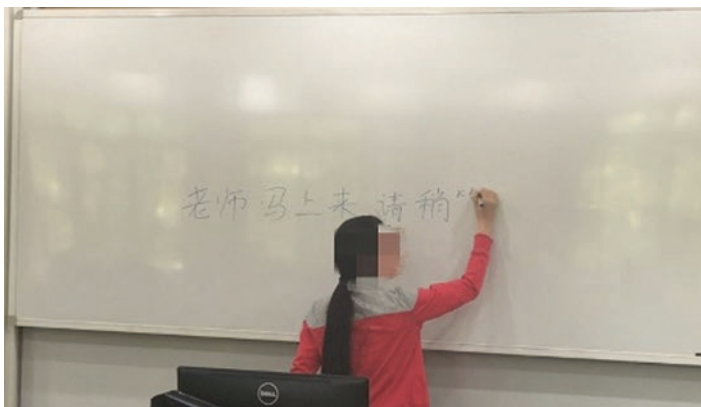


Fig. 3 “The professor will come soon. Please wait” in Chinese on whiteboard

6 Discussion

Language ideologies may be multiple within one cultural setting (Woolard, 2021), and our findings in this chapter reveal the tensions and paradoxes between different language ideologies within this study abroad program in China. The institutional choice of English demonstrates that, though China’s policies sometimes promote the Chinese language as a tool to enhance China’s soft power abroad (Hubbert, 2019; Liu & Tao, 2012), in other instances, language is seen as an obstacle toward increasing China’s cultural and political influence. As revealed in our study, English is institutionally perceived as a more efficient tool to increase the enrollment of international students and to achieve the educational goals of the BRI. However, the findings also highlight how the very choice of English could create the opposite impression among the students, that the program was promoting Westernization rather than the Chinese experience or culture.

The frequent English usage in this study was also related to the institutional arrangement that separated domestic and international students, which helped reproduce the identity category of “foreign students” in China. The students actively participated in the ideological distinction between themselves and the local Chinese students, the latter of whom were assumed to be speakers of Chinese and simultaneously not proficient in English. Meanwhile, the students in the program perceived Anglophonization of their program to mean that it was more Americanized and themselves more competitive on the global job market that appeared to prioritize English (Heller & McElhinny, 2017; Park & Bae, 2009). However, in parts of East Asia college admission without taking the entrance exam in the local language could be viewed as an easier alternative for students from overseas, who by extension were also deemed as less well trained in subjects such as math and sciences (Kanno, 2003). The faculty and staff thus considered these international students as being less qualified academically than their local Chinese peers, and the use of English in the program to be an academic compromise.

The social imaginary of the study abroad context often takes boundary crossing for granted (Kubota, 2016). However, in the transnational context of the program, we observed multiple boundaries being (re)produced and (re)negotiated by the program and by the students themselves. The students and faculty members routinely made comments regarding citizenship and language use in relation to business, ideologically framing and sometimes contesting these national boundaries. The institutional display of the flags and the confusion it caused among the students were also evidence of how the program and the students, respectively, made sense of “international” in a world bounded by geographical borders and passports. Our findings shed light on the contradictions between different webs of ideologies, both related to language use and beyond. Their (re)negotiations of linguistic, citizenship, and cultural boundaries can contribute to an improved understanding of boundary crossing in study abroad and provide implications for researchers and administrators interested in language and international education.

Meanwhile, the oft-quoted metaphor of a “foreign bubble” among the students was used to both define and justify the exclusive use of English among themselves. Yet such a monolingual expectation about study abroad becomes contested (Diao & Trentman, 2021; Surtees, 2016). While these international students were assumed to be unable to attain enough fluency in Chinese to do academic or professional work, they were also simultaneously expected to know enough Chinese to navigate through various academic situations that involved extensive use of Chinese or frequent code-switching. Although the program was marketed as an academic experience in English only, in actuality it expected its students to be functional bilinguals who can readily cross linguistic boundaries between Chinese and English.

The BRI is often viewed as one of the ways that China is rivaling the dominance of the U.S. on the global stage, and many of the worlds’ most developed Anglophone countries are notably absent from it. By focusing on the international educational exchange situated in China during the BRI, the current study serves as a departure from the traditional focus in the research literature on the developed world and join the emerging scholarship that explores study abroad within the Global South (e.g., Thomas, 2021). As revealed in our findings, oftentimes neither the destination nor the students’ home countries are Anglophone in this context. Our results thus not only problematize the assumption that English should be used as a default in global educational migration; they also raise questions regarding China’s growing influence on language use and education in the developing world and urge future research to further cross these linguistic and social boundaries.

Finally, as we write this chapter, programs intended for international students in China are facing unprecedented challenges. China’s strict travel restrictions that were put in place to prevent a resurgence of the COVID-19 pandemic have meant that many programs, like the one we focused on, that were intended for international students have practically halted their recruitment efforts. This new context may soon lead to another ideological shift in conceptualizing international education and its relationship to the BRI. Our chapter is only intended to capture the moment when the data were generated, and there need to be more conversations regarding how to sustain opportunities for both study abroad and research about language learning/use during study abroad.

References

- Billig, M. (1995). *Banal nationalism*. SAGE.
- Blommaert, J. (1999). *Language ideological debates*. De Gruyter Mouton.
- Chen, B. (2018). *Press conference*. Retrieved from http://www.chisa.edu.cn/rmtnews1/ssyl/201803/t20180316_3023530.html
- Coleman, J. A. (2006). English-medium teaching in European higher education. *Language Teaching*, 39(1), 1–14.
- De Costa, P. I. (2016). *The power of identity and ideology in language learning: Designer immigrants learning English in Singapore*. Springer.

- Diao, W. (2021). Language use, class and study abroad in China. In W. Diao & E. Trentman (Eds.), *Language learning in study abroad* (pp. 121–144). Multilingual Matters.
- Diao, W., & Trentman, E. (2016). Politicizing study abroad: Learning Arabic in Egypt and Mandarin in China. *L2 Journal*, 8(2), 31–50.
- Diao, W., & Trentman, E. (Eds.). (2021). *Language learning in study abroad: The multilingual turn*. Multilingual Matters.
- Donato, R., & Tucker, G. R. (2007). K-12 language learning and foreign language educational policy: A school-based perspective. *Modern Language Journal*, 91(2), 256–258.
- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.
- Dong, H. L. (2001). Duanzao guoji fuhexing rencai. *People's Daily*. Retrieved from <https://rnrbrb.online/simple/?t1283966.html>
- Douglas Fir Group. (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, 100(S1), 19–47.
- Gao, X. (2006). Understanding changes in Chinese students' uses of learning strategies in China and Britain: A socio-cultural re-interpretation. *System*, 34(1), 55–67.
- Gee, J. P. (2004). *An introduction to discourse analysis: Theory and method*. Routledge.
- Gere, L. (2017). *Maps of the silk road*. Retrieved from <http://beltandroadcenter.org/2017/12/15/maps-of-the-silkroad/>
- Gil, J., & Adamson, B. (2011). The English language in Mainland China: A sociolinguistic profile. In A. Feng (Ed.), *English language education across greater China* (pp. 23–45). Multilingual Matters.
- Gore, J. (2005). *Dominant beliefs and alternative voices: Discourse, belief, and gender in American study abroad*. Routledge.
- Guardado, M. (2018). *Discourse, ideology and heritage language socialization*. De Gruyter Mouton.
- Hadley, G. (2017). *Grounded theory in applied linguistics research: A practical guide*. Routledge.
- Hadley, G. (2019). Grounded theory method. In J. McKinsley & H. Rose (Eds.), *The Routledge handbook of research methods* (pp. 264–276). Routledge.
- Heller, M., & McElhinny, B. (2017). *Language, capitalism, colonialism: Toward a critical history*. University of Toronto Press.
- Hubbert, J. (2019). *China in the world: An anthropology of Confucius institutes, soft power, and globalization*. University of Hawaii Press.
- Irvine, J. T. (2012). Keeping ethnography in the study of communication. *Language et Societe*, 1, 47–66.
- Kanno, Y. (2003). *Negotiating bilingual and bicultural identities: Japanese returnees betwixt two worlds*. Routledge.
- Kinginger, C. (2008). Language learning in study abroad: Case studies of Americans in France. *The Modern Language Journal*, 92(s1), 1–124.
- Kubota, R. (2008). Critical approaches to teaching Japanese and culture. In J. Mori & A. S. Ohta (Eds.), *Japanese applied linguistics: Discourse and social perspectives* (pp. 327–352). Continuum.
- Kubota, R. (2013). ‘Language is only a tool’: Japanese expatriates working in China and implications for language teaching. *Multilingual Education*, 3(1), 4. Retrieved from <https://multilingual-education.springeropen.com/track/pdf/10.1186/2191-5059-3-4.pdf>
- Kubota, R. (2016). The social imaginary of study abroad: Complexities and contradictions. *The Language Learning Journal*, 44(3), 347–357.
- Li, S., & Tucker, G. R. (2013). A survey of the U.S. Confucius institutes: Opportunities and challenges in promoting Chinese language and culture education. *Journal of the Chinese Language Teachers Association*, 48(1), 29–53.
- Liu, W. (Ed.). (2018). *China's belt and road initiatives: Economic geography reformation*. Springer.
- Liu, J., & Tao, H. (2012). Negotiating linguistic identities under globalization: Language use in contemporary China. In J. Liu & H. Tao (Eds.), *Chinese under globalization: Emerging trends in language use in China* (pp. 203–212). World Scientific.

- McGroarty, M. (2002). Evolving influence on educational language policies. In J. W. Tollefson (Ed.), *Language policies in education* (pp. 17–36). Lawrence Erlbaum.
- Ministry of Education of the People's Republic of China. (2017). Retrieved from http://www.moe.gov.cn/jyb_xwfb/gzdt/s5987/201803/t20180329_331771.html
- Ministry of Education of the People's Republic of China. (2000–2018). 年度出国留学人员情况统计. Retrieved from <http://www.moe.gov.cn/>
- National Development and Reform Commission (NDRC), Ministry of Commerce (MOC), & Ministry of Foreign Affairs (MOFA) of the People's Republic of China. (2015). *Vision and actions on jointly building the Belt and Road*. Retrieved from http://de.china-embassy.org/det/zt/yidaiyilude/201503/t20150330_3126178.htm
- Pan, L., & Seargeant, P. (2012). Is English a threat to Chinese language and culture? *English Today*, 28(3), 60–66.
- Park, J. S. Y., & Bae, S. (2009). Language ideologies in educational migration: Korean *jogi yuhak* families in Singapore. *Linguistics and Education*, 20(4), 366–377.
- Pérez-Milans, M. (2012). Beyond “safe-talk”: Institutionalization and agency in China's English language education. *Linguistics and Education*, 23(1), 62–76.
- Robertson, S. L., & Kedzierski, M. (2016). On the move: Globalising higher education in Europe and beyond. *The Language Learning Journal*, 44(3), 276–291.
- Rose, H., & McKinley, J. (2018). Japan's English-medium instruction initiatives and the globalization of higher education. *Higher Education*, 75(1), 111–129.
- Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). SAGE.
- Silverstein, M. (1979). Language structure and linguistic ideology. In P. R. Clyne, W. F. Hanks, & C. L. Hofbauer (Eds.), *The elements: A parasesion on linguistic units and levels* (pp. 193–247). Chicago Linguistic Society.
- Smith, K. (2004). Studying in an additional language—What is gained, what is lost, and what is assessed? In R. Wilkinson (Ed.), *Integrating content and language: Meeting the challenge of a multilingual higher education* (pp. 78–93). Universitaire Pers Maastricht.
- Song, J. (2010). Language ideology and identity in transnational space: Globalization, migration, and bilingualism among Korean families in the USA. *International Journal of Bilingual Education and Bilingualism*, 13(1), 23–42.
- Surtees, V. (2016). Beliefs about language learning in study abroad: Advocating for a language ideology approach. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 27, 85–103.
- Thomas, J. A. (2021). Ghanaian multilinguals on study abroad in Tanzania: Learning Swahili through Akan/Twi and cultures of storytelling. In W. Diao & E. Trentman (Eds.), *Language learning in study abroad* (pp. 13–42). Multilingual Matters.
- Trentman, E., & Diao, W. (2017). The American gaze east: Discourses and destinations of U.S. study abroad. *Study Abroad Research in Second Language Acquisition and International Education*, 2(2), 175–205.
- Tucker, G. R. (1968). Judging personality from language usage: A Filipino example. *Philippine Sociological Review*, 16(1/2), 30–39.
- Tucker, G. R., & Donato, R. (2003). Implementing a district-wide foreign language program: A case study of acquisition planning and curricular innovation. In D. Tannen & J. E. Alatis (Eds.), *Georgetown University round table on languages and linguistics 2001* (pp. 178–195). Georgetown University Press.
- Wang, Y. (2020). *Studying in China but in an English program: Language ideologies among study abroad students in the age of Belt and Road*. Unpublished Doctoral Dissertation. University of Arizona.
- Wodak, R. (2015). Critical discourse analysis, discourse-historical approach. In K. Tracy (Ed.), *The international encyclopedia of language and social interaction*. Wiley.
- Woolard, K. A. (2021). Language ideologies. In J. M. Stanlaw (Ed.), *The international encyclopedia of linguistic anthropology* (pp. 1–21). Wiley.
- Xi, J. P. (2017, October 26). 在中国共产党第十九次全国代表大会上的报告. *People's Daily Online*. Retrieved from <http://english.mofcom.gov.cn/article/pressconferencehomepage/foreigninvestment/201711/20171102667894.shtml>

When Transfer Transfers: Applying Cross-Linguistic Reading Transfer Theory to Language of Instruction Policies in Low- and Middle-Income Countries



Pooja R. Nakamura

Abstract In low- and middle-income countries (LMICs) across the world, students are learning to read in languages they do not use or understand, and then transitioning to a new language of instruction during primary education. This disparity between home and instructional languages coupled with a demand for instruction in post-colonial and international languages contributes to the fact that more than half of the world’s LMIC children are unable to read in any language by the time they are 10 years old (World Bank, Loud and clear: effective language of instruction policies for learning. The World Bank, Washington, 2021a). The theoretical notions of cross-linguistic transfer in biliteracy acquisition hold unique promise in addressing these language of instruction policy questions by lending evidence for what languages students should be taught in “first,” and how to structure effective transition to a new language of instruction. The importance of stakeholder engagement and boundary crossing for application of theory to policy and practice sustainability in low-resourced LMIC education contexts is also discussed.

Keywords Transfer · Biliteracy · Education policy

1 Introduction: The First Boundary

For decades, two streams have flowed in seeming parallel: the increase in theoretical knowledge and empirical evidence on how young children learn to read two (or more) languages, and the increase in the number of children in bilingual and multilingual contexts in low- and middle-income countries (LMICs) who are simply not learning to read (in any language) – a state of “reading poverty” (UNESCO Institute for Statistics, 2017; World Bank, 2017). For the most part, the first stream has been

P. R. Nakamura (✉)
American Institutes for Research, Arlington, VA, USA
e-mail: pnakamura@air.org

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

D. Zhang, R. T. Miller (eds.), *Crossing Boundaries in Researching, Understanding, and Improving Language Education*, Educational Linguistics 58, https://doi.org/10.1007/978-3-031-24078-2_16

341

the concern of academics and scholars and the second stream has been the concern of policy makers and program implementers, and the boundaries between them have remained relatively firm.

In this chapter, I argue that one of the most promising ways to alleviate reading poverty is to make the boundaries between theory (especially reading transfer theory) and policy (especially language of instruction [LoI] policy) and practice (especially biliteracy teaching practice) more intentionally permeable, with a focus on LMICs. Specifically, by drawing on cross-linguistic transfer theories of reading, two main policy areas are discussed through a global lens (Tucker, 1998): (1) determining the initial language of literacy to begin instruction to build foundations for transfer; and (2) timing of transition of instruction from one language to another, with empirical thresholds for transfer. In addition, I discuss the possibility of expanding the theoretical underpinnings of reading transfer to incorporate threshold mechanisms as well as the importance of stakeholder engagement (Donato & Tucker, 2010) for application of theory to improving literacy outcomes in low-resourced LMIC education contexts.

2 Problem Statement: Reading Poverty

As the international education community transitioned from the Education for All initiatives (World Bank, 2014) and the Millennium Development Goals (MDGs) (United Nations, 2015) to the Sustainable Development Goals (United Nations, 2022) focus shifted from education access to quality of learning. This shift was prompted by the fact that even though an unprecedented number of children were in schools (prior to the onset of the global COVID pandemic in early 2020), many were simply not learning. Around 6 out of 10 (a total of 617 million) children and adolescents globally are not able to read with minimum proficiency (UNESCO Institute for Statistics, 2017). The World Bank revealed that more than half of all children in LMICs are unable to read a simple story by the time they are 10 years old (World Bank, 2017). The COVID pandemic is likely to accelerate that number to about 70% of all children in LMICs in the coming years (World Bank, 2021b).

Although there are multiple reasons for this state of reading poverty, a critical one is language of education. Approximately 37% of children across LMICs are being taught to read for the first time in school in languages they do not use or understand (World Bank, 2021a). In some countries, up to 90% of primary school children have a mismatch between the language(s) they speak in their home and the language(s) in which they are taught in school. Within countries, children in the lowest socioeconomic bracket are less likely to have access to education in a language they use and understand, and furthermore lack the family and home resources to overcome this language mismatch.

Beyond the initial LoI mismatch issue, even in countries and contexts where children are initially taught in a language they have proficiency in, they are often required to transition to a new LoI relatively quickly. Countries (and regions within

countries) continually change their LoI policies, including the LoI transition year. For example, Rwanda's LoI policy changed from French to English in 2008; from English to Kinyarwanda in 2015; and from Kinyarwanda back to English in 2019 (Edwards, 2019). Mozambique's Ministry of Education is poised to shift from a mostly Portuguese-only LoI policy to a more gradual transition model starting with only oral Portuguese skills (taught as a subject) in Grades 1 and 2, followed by a variety of subjects taught in Portuguese from Grade 4 through 6, followed by a full consolidation of transition of LoI in Grade 6 (MINEDH, Government of Mozambique). In India, the regions and socioeconomic brackets of the school tend to determine the LoI transition grades and timings, and there is wide variety across regions and across time periods (Annamalai, 2005).

These constantly shifting LoI policies and practices reflect the fact that there is limited evidence on effective LoI transition timing, and also that LoI policies are highly politically charged in many countries. It is clear that students will learn to read with comprehension only in languages they speak and understand, while it is also clear that there is an ever-increasing demand from communities for education in a post-colonial or international language, given its association with socioeconomic mobility (Azam et al., 2013; Chakraborty & Bakshi, 2016; Coleman, 2011). This leads to the dilemma of how to balance the need for longer "mother tongue" or local language LoI with the demand for earlier introduction of the post-colonial LoI.

One way to offer empirical reconciliation to this macro policy question is to examine the cognitive question of whether – and if so, when – there is a point at which a child is cognitively and linguistically "ready" to be introduced to literacy instruction in a second (or later acquired) language. In other words, would it be reasonable to expect that there is a precise, measurable level of mastery of first language (L1) literacy skills at which cross-linguistic transfer is triggered and below which transfer is much less likely? The answer to this question could provide important information on how to structure effective biliteracy instruction programs and policies.

3 Stakeholders: Why Does Biliteracy Acquisition Matter to Them? And Why Do They Matter?

Reading skills are one of the most critical foundational skills for academic achievement and later career development. Students' foundational literacy and numeracy performance on international assessments explains more than 70% of the variation in a nation's economic growth (Hanushek & Woessmann, 2008). Several bilateral and multilateral donor agencies providing education funding to LMICs have prioritized foundational literacy in their agendas (UNESCO, 1990, 2000; USAID, 2019; World Bank, 2021a), which in turn has led country governments to also prioritize foundational literacy in their policies (e.g., India's National Initiative for Proficiency in Reading with Understanding and Numeracy Bharat Initiative [NIPUN Bharat];

Ministry of Education, India, 2021). At the same time, in a majority of LMICs in which the learning crisis is most profound, students are learning in bilingual or multilingual contexts. This bilingualism or multilingualism plays out in a variety of different ways, including multiple languages being spoken in each community or in each home, or even cases when only one language may be spoken in the home and community, but students are required to learn another language in school from the beginning of pre-primary or formal primary school instruction.

This combination of widespread bilingualism and multilingualism and a heavy focus on development of foundational literacy skills leads to a series of open empirical questions for education stakeholders in LMICs. Ministries of education, teacher training institutions, curriculum developers, and parents and community members are all keenly interested in effective ways to design LoI and biliteracy programs and policies that require the least number of resources. Critically, there is a deep need for stakeholder buy-in and ownership of the evidence on how to improve literacy programing in multilingual LMIC contexts.

While there has been an increase in the evidence base for effective LoI policies in LMICs (e.g., Evans & Mendez Acosta, 2021; Nag et al., 2019), the “brokering” of that knowledge between the evidence-generators and the evidence-users requires careful planning, including contextual analysis, definition of uptake objectives across keys stakeholders, mapping of the audience, creation of a detailed communications plan, and monitoring of research uptake (Brown et al., 2018). In LMICs, several programs have shown that stakeholder buy-in and ownership of the early learning program is a necessary component of effectiveness and scalability. For example, Beeharry (2021) argued that for the attainment of the SDG goal 4 agenda of foundational literacy and numeracy, international stakeholders (bilateral donor agencies), governments, civil society organizations, and nongovernmental organizations all not only need to buy in to the importance of improving literacy outcomes, but also need to be held accountable for achieving it. Brunett (2019) echoes the importance of stakeholder buy-in by calling for discussion, funding, and a combination of technical and financial support to bolster early learning outcomes in LMICs. In a powerful example, Bold et al. (2013) demonstrate that small-scale randomized controlled trials show significant impact on early learning; however, the same programs show no impact when the government agencies or other constituents who run them have limited buy-in and ownership. Piper (2017) also stresses that when research is conducted without considering government buy-in and the political economy within which it will be applied, it risks being distorted, ineffective at scale, or simply ignored.

In the United States, Donato and Tucker (2010) also stress the ultimate role of the constituents outside the daily working of a foreign language program in ensuring its sustainability. Parents and community members, school board members, other subject area teachers, and school district-university partnerships are all examples of stakeholders in the West who ultimately contribute to the culture of a sustainable foreign language program.

Moving across geographic boundaries, recent interviews were conducted with a variety of stakeholders on LoI transition policies and biliteracy (or multi-literacy)

development in Côte d'Ivoire, Democratic Republic of Congo, and Senegal (American Institutes for Research, 2022). The interviewees included stakeholders from each country's Ministry of Education, international organizations that are implementing large-scale early education programs in these countries, donor organizations (USAID and UNESCO), and academics and research organization affiliates. Every stakeholder respondent noted that they were interested in evidence to support the question of when the best time is to transition from one language of instruction to another. Through these interviews and other stakeholder dissemination workshops conducted in other projects in India (Nakamura & de Hoop, 2014) and Ethiopia (Nakamura et al., 2019a, b), for example, the stakeholders invariably point to the need to find out the timing of transition for effective bilingual and multilingual education policies. Whether it is a grade at which transition is most likely to be effective or a skill set that needs to be acquired for a child to be sufficiently ready to learn to read in a new language, there appears to be a major gap in knowledge in determining *how much is good enough* to introduce reading in a new language – and how to convert that knowledge into LoI policy.

These questions – borne out of boundary crossing between stakeholders' macro-level priorities and academics' concern with the micro-processes in cross-linguistic reading transfer theory – lead to the following areas of inquiry: how we determine the first language in which to introduce literacy instruction in multilingual contexts, and how we determine a level of readiness in particular reading sub-skills to determine when a child has sufficiency in L1 reading to develop L2 reading skills. I will examine these two areas in the following sections.

4 The “Initial” Language of Instruction

The selection of an initial language of instruction – that is, the first language a child will learn to read in – has very limited empirical basis in many LMIC educational contexts. Theoretically, reading models suggest that, across language and orthography types, learning to read with comprehension requires oral language proficiency. For instance, the Cognitive Foundations for Reading Acquisition model, Hoover and Tunmer's (2020) updated version of the Simple View of Reading (Hoover & Gough, 1990), suggests that language comprehension – comprised of background knowledge, linguistic knowledge, phonological knowledge, syntactic knowledge, and semantic knowledge sub-skills – is one of two major pillars (the other being automatic word recognition) required for developing reading comprehension ability. Perfetti's (2003) Universal Grammar of Reading puts forward that reading ability is, essentially, the link between spoken language and written language. In other words, if a child cannot map the written symbols to their phonological and morphological representations (Perfetti, 2003), and then to a larger inference making system and background knowledge to make sense of larger swathes of print (Perfetti & Dunlap, 2008), reading comprehension becomes unattainable. August and Shanahan (2006) review the links between oral language proficiency and reading outcomes

from the word level up to the discourse level. Nag and colleagues demonstrate, through two systematic reviews, that the inter-related concepts of oral language skills and home-school language mismatch are consistently strong predictors of early literacy sub-skill attainment (Nag et al., 2014, 2019). These studies extend what we know about the negative impacts of home-school language mismatch from OECD countries (e.g., Karlsen et al., 2017; Scheele et al., 2010) to LMICs.

Despite this theoretical and empirical basis for beginning education in a language (or languages) that a child understands, in most LMIC educational settings, there are high levels of mismatch between the language children speak at home and the one they learn initially in school (Brock-Utne, 2021; Heugh, 2012). Since 1953, the United Nations Educational, Scientific and Cultural Organization (UNESCO) has advocated for mother tongue-based multilingual education and attempted to dispel several myths surrounding its ineffectiveness (UNESCO, 1953). The myths include, for example, the beliefs and arguments that multilingualism is an impediment to unity, there are too many languages to teach in schools, parents and teachers have no demand or use for the mother tongue within the educational systems, if too much time is spent learning the mother tongue it will be at the cost of learning the international language, etc. In 2010, the United Nations developed a comprehensive policy brief on how to implement mother tongue programs in Africa (Ouane & Glanz, 2010), based on research linking mother tongue-based multilingual education (MTB-MLE) with stronger educational outcomes in both the mother tongue and in the second (later) acquired language (Alidou et al., 2006; Heugh et al., 2007; Ouane & Glanz, 2010), as well as research showing the costs and benefits in the long term (Heugh, 2012). In the United States, Collier and Thomas (2017) present powerful data from a 32-year longitudinal study to reveal that it takes an average of 6 years of high-quality dual language instruction (with at least 50% of the instruction occurring in the L1) in order for English learners to close the achievement gap between bilingual and monolingual children. Together, these studies make a consistent case for initial instruction in languages children speak and understand.

A complicating factor that needs to be considered for making the initial LoI choice is how the manifestations of multilingualism – and the values that each language carries for societies, communities, and individuals – vary across boundaries. The configurations of the languages chosen for education by decision makers and those driving education demand in communities reflect links to identity and heritage cultural value versus socio-economic and transactional benefits. Olshtain and Nissim-Amitai (2004) distinguish between “natural multilingualism,” as is common in countries like Ethiopia, Kenya, and India, and “multilingualism resulting from immigration,” which is common in the Netherlands, the United States, and Israel, for example. In OECD countries, multilingual education is framed as a virtuous enhancer or an economic and cultural “accessory” for culturally adept citizens (Kohler, 2017). Bialystok (2018) also makes a critical distinction between “bilingual education” and “education of bilingual children.” At its heart, this distinction is whether the focus of the education system and goals is robust bilingual education outcomes (i.e., higher education outcomes in both languages), or whether the focus is on the native-language supports and scaffolds that need to be in place only for the

purpose of being able to successfully transition from that native language into higher education in the second (or later) language. This distinction in framing has critical implications for how likely the policy will be taken up in a sustainable fashion, as multilingualism can be quickly shunned for a preference for “unity” or socio-economic mobility – both perceived and real.

Therefore, when developing an initial LoI policy, decision makers are faced with a point on a spectrum from a nearly 100% match between the LoI and the language spoken at home by all students to an almost 100% mismatch wherein all students speak a language at home that does not match the language that policy makers want students to use in schools. Spolsky and Shohamy (2000) lay out a range of alternatives within these contexts, understanding that most countries fall somewhere on this spectrum. In cases of mismatch and multilingualism, countries could opt for so-called “submersion,” in which all students are educated in the language of the country, which has yielded relatively lower academic achievement scores compared to other options. Alternatively, countries could choose a dual language or multilingual education program, in which students learn some combination of their home language and school language throughout their education, leading to stronger educational outcomes in both the L1 and L2. In many LMICs, the model is transitive, starting with one language and then transitioning to the national/post-colonial, and in OECD countries, the model is more additive, without replacement of the LoI with a new one.

LMICs’ educational contexts are also characterized by substantial differences between urban and rural communities in terms of degree and type of multilingualism. About 25% of the world’s population will live in urban poor “slum” communities by 2030 (Habitat for Humanity, 2022). These communities across the world are characterized by complex multilingualism. In one instance in Bangalore, India, students in slum communities use of three to five languages daily for different purposes and code-switch through the day (Reddy, 2011). Yet, little empirical evidence exists on students’ level of linguistic competence in the multiple languages that are spoken by these children, which has significant implications for choice of the initial LoI.

In more linguistically homogenous, rural schools in LMICs, the problem plays out differently. Although there may be only one (or one predominant) language spoken in the community, the policy does not allow for that language to be used in schools, essentially barring children from learning to read first in a language in which they have any oral language proficiency. In a language mapping study (Nakamura et al., 2017) in two provinces in Mozambique, Nampula and Zambezia, children were asked to look at a picture with a detailed scene as a prompt and list all the words that came to their mind. This “semantic fluency” test was conducted in all languages (separately) that the child claimed that they could speak and understand even to a very small degree. In this way, it was possible to establish language dominance and oral language skills and utilize that information to determine initial languages of instruction. The first key finding from this study was that there were roughly normal distributions when children were tested in a range of languages and when the test focused on oral language skills. A second key finding was that there

were large mismatches between census (self-reported) data and objectively measured language proficiency data, underscoring the fact that census (or other self-reported) data may be misleading sources of information for policy makers to use to determine initial LoI policy. Third, the study also underscored that approximately 62% of the schools were linguistically heterogenous, defined as having at least two different sets of languages spoken by students in the classroom. Most remarkably, at the time of data collection, 73% of the students in the sample did not have oral language proficiency in the languages in the initial LoI of the schools (Portuguese).¹

In sum, it is clear that the selection of the first language a child will be taught *in* is a challenging question for education stakeholders in LMICs fraught with issues around linguistic competence levels, identity, political will, and socioeconomic access and mobility. As long as there is a distinction between “bilingual education” and “education for the bilingual child,” stakeholders will not be able to develop a program that embraces the language of the child’s home as the first language of education. It is also evident that there is limited data and evidence upon which to make the decision of which should be the first language in many areas, which in turn makes a shaky foundation for transitioning to learning to read in a new language.

5 The “Next” Language: Cross-Linguistic Transfer and Language Transitioning

In most LMIC educational contexts, students need to transition from one LoI to another at some point in their primary or secondary schooling (if they began education in a local language at all, which, as noted above, is not the case in many contexts). While the evidence has continued to build for beginning learning to read (and learning in general) in a language that the child has strong oral language proficiency in, there is still very limited evidence on how and when to make an effective transition to a new LoI – most often the post-colonial language (e.g., English in Kenya and India, French in Senegal and Democratic Republic of Congo, Portuguese in Angola and Mozambique). In several countries there may be a third language – either as a LoI or as a subject (e.g., Hindi in India in regions where Hindi is not a “mother tongue” or regional language, or Kiswahili in South Africa and Kenya in regions where Kiswahili is not a “mother tongue” or regional language).

While this LoI transition entails a set of skills and knowledge that go beyond biliteracy acquisition, the ability to learn to read – and then read to learn – in two languages is central to LoI transitioning. Biliteracy acquisition, in turn, is characterized in significant ways by cross-linguistic relationships. As such, it is

¹ The policy has since shifted in these areas to mother tongue instruction in Elomwe, Echwabo, and Emakhuwa, and there is an expectation that a new bilingual education policy will be rolled out in 2022, in which children will begin learning Portuguese as a second language in Grade 1 and gradually transition, subject-by-subject, from Grade 4 through Grade 6, to a full Portuguese-only LoI policy in Grade 6 (MINEDH, Government of Mozambique).

reasonable to argue that a deeper understanding of how cross-linguistic relationships play out in a variety of settings in LMICs should play a critical role in determining transition sequencing, timing, and planning. In essence, the question most critically needing answering for policy decision makers in LMICs is whether there is a point at which a child is most likely to be “ready” to transfer their skills from an L1 to an L2 (or vice versa), and if so, when that is. It is through the lens of this question that I briefly describe how L1-L2 transfer and cross-linguistic relationships have been studied and conceptualized thus far.

One of the most common ways cross-linguistic relationships have been conceptualized in the literature is as “transfer” of skills from one language to another. Over the last few decades, the defining mechanisms and underlying processes underpinning transfer have shifted and evolved to into a complex, multi-dimensional, and dynamic process. In one of the earliest theoretical formulations of transfer, Lado (1957) focused on the interfering effects of the L1 linguistic and orthographic properties on processing L2. Subsequently, Cummins (1979, 1981) proposed the notion of a common underlying proficiency within the linguistic interdependence hypothesis, in which he contended that acquisition of L2 reading ability – especially academic content – is contingent upon the degree to which a child has sufficiently mastered context-independent, cognitively demanding skills and knowledge in their L1. At the same time, he also proposed the linguistic threshold hypothesis, in which he contended there is a specific level of L1 or bilingualism that is needed for the positive impact to be manifested in bilingual reading. Herein, the seed is laid that there is some level of “sufficiency” for the L1 reading ability to be able to trigger the transfer process. However, what this level is, and precisely what constitutes these transferrable skills and processes remained unspecified.

Extending this notion, several studies have shown that there are significant correlations between corresponding skills in the L1 and L2, including for example, phonological awareness (Branum-Martin et al., 2006; Comeau et al., 1999; Verhoeven, 1994; Wawire & Kim, 2018), decoding (da Fontoura & Siegel, 1995; Geva et al., 1997; Gholamain & Geva, 1999), and working memory and rapid naming (Abu-Rabia & Siegel, 2002; Geva & Siegel, 2000; Gholamain & Geva, 1999). While this set of studies can be interpreted as evidence for the transfer of L1 skills to their corresponding L2 skills, most of the authors argue for a “common underlying proficiency” that explains the relationship. The flip side of this same body of research is that there are also orthographic properties that constrain the relationship between parallel sub-skills in two languages (Geva & Siegel, 2000; Kim, 2009; Reddy & Koda, 2013; Wang et al., 2006, 2009). These findings for both a common underlying proficiency and script-dependent transfer are also reflected in trilingual literacy development (Abu-Rabia & Siegel, 2002; Rickard Liow & Poon, 1998) and have been validated through reviews and meta-analyses (e.g., August & Shanahan, 2006; Melby-Lervåg & Lervåg, 2014).

The Transfer Facilitation Model (TFM) focuses on metalinguistic awareness – especially phonological awareness and morphological awareness – as the key component of transfer (Koda, 2008). The TFM highlights that once the language-neutral components of metalinguistic awareness (as well as higher-order, language-neutral

components such as sound-to-symbol mappings, symbol-to-meaning mappings, and inference making, for example) reach a level of automaticity, they move from being sharable cognitive and linguistic resources to shared cognitive and linguistic resources. In other words, these resources are automatically drawn upon to facilitate L2 reading. While L2 proficiency, L2 input properties, as well as L1-L2 distance continue to moderate this cross-linguistic resource sharing, critical to our discussion is the conceptual notion that there is a point at which there is “sufficient” automaticity in order for the resources to shift from being sharable to being shared. Many studies have provided empirical evidence for these theoretical foundations (for meta-analyses and reviews, see Ke et al., 2021; Koda & Reddy, 2008; Melby-Lervåg & Lervåg, 2014).

Recently, it has been proposed that transfer be couched within a larger macro framework of interactive factors (Chung et al., 2019). In addition to the cognitive, metacognitive, orthographic, and linguistic factors that determine transfer and resource-sharing processes (such as L1-L2 distance, L1-L2 proficiency levels, and language complexities), it is also imperative to consider educational setting and research methodology.

In LMICs, types of bilingual and multilingual contexts are varied. As discussed above, urban poor communities tend to have more linguistically heterogenous contexts and so-called *functional multilingualism* wherein students speak more than 3 languages for varying functions (e.g., school vs. home vs. the community market) with varying degrees of proficiency, while rural areas tend to have one local language primarily spoken. It is in these urban areas that the question of bidirectionality of transfer and cross-linguistic relationships come up. Most studies discussed above looked at L1 transfer after L1 skills are relatively well established. Studies have shown limited transfer from an L2 to an L1 (Gottardo et al., 2014); however, when the acquisition of two languages is simultaneous or near-simultaneous, research points to a more bidirectional relationship between the two languages (Deacon et al., 2009; Kim & Piper, 2019; Pretorius & Mampuru, 2007; Wagner et al., 1994).

Taken together, studies on biliteracy acquisition have focused on cross-linguistic relationships between languages and the role of metalinguistic awareness in predicting within- and across-language reading outcomes. The cross-language relationships are relatively robust in terms of language-neutral and sharable constructs such as phonological awareness and decoding, but less strong in language-specific factors such as oral vocabulary (Melby-Lervåg & Lervåg, 2011). This leaves policy makers in several LMIC contexts without information on precisely when to make the decision to introduce literacy instruction in a new language in their varied contexts – a question I turn to next.

6 Transfer Thresholds

Given that academic research sometimes proceeds parallel to – rather than in conjunction with – conversations with LMIC stakeholders, the focus tends to be on forging the theoretical frontiers of biliteracy acquisition. Such boundaries between practice and theory are common across disciplines and even within the sphere of bilingual reading (Grabe, 2009). What is most interesting here, however, is that this theoretical frontier is, in fact, one of the most critical gaps in knowledge for stakeholders in designing bilingual and biliteracy programs, leading to a prime opportunity for boundary crossing to lead to significant educational impact in LMICs. The practical question of “when is the right time to introduce literacy acquisition in a second (or later acquired) language?” translates exactly into the theoretical question of “is there a point at which children are ready to transfer their L1 skills to L2 literacy learning?”

The idea that there may be a tipping point of readiness for transfer is not new nor is it developed in a theoretical vacuum. Scholars of English reading have made a distinction between those reading skills that are constrained or within a finite learning boundary and those that are unconstrained or within a more boundaryless space in terms of the number that need to be acquired (NICHD, 2000; Snow & Mathews, 2016). The former refers to skills that are finite and can be considered “fully” acquired, such as the names of letters or spelling rules. The latter refers to skills that are relatively infinite and no one person can be considered to know them all, such as vocabulary or background knowledge. Kim et al. (2020) extend this notion to reading in several languages in LMICs.

At the same time, biliteracy research has pointed to the notion of “threshold” level mastery for transfer to be triggered. For example, the main thrust of Cummins’ (1979) linguistic threshold hypothesis is that reading skills need to attain a sufficient level of mastery for transfer to occur. Koda’s (2008) TFM also makes the case that L1 skills are shared only after they reach a degree of automaticity. The implication of these two theoretical lines is that there may be a non-linear relationship between L1 and L2 skills in the constrained skills transfer. In other words, constrained skills like symbol-sound correspondences are likely to be transferred when they reach a point of mastery in the L1; whereas the unconstrained skills – while supportive to L2 reading acquisition – may not necessarily be conducive to identifying a specific point or threshold level of mastery.

Most studies that have provided evidence for transfer mechanisms between corresponding sub-skills utilize linear regression analyses. While this method has shed considerable light on which L1 and L2 sub-skills are related, it does not answer the question of the degree of mastery of L1 sub-skills needed for transfer to L2 reading to begin kicking in. In other words, regressions reveal that with the increase of one point in one skill (most likely in L1) there is a corresponding point increase in another skill (most likely in L2); however, we cannot determine whether the degree of increase in L2 (i.e., the slope of the linear relationship) changes at any particular point – or points – along the acquisition of the corresponding transferrable L1 skill.

To address this gap and answer the questions posed by stakeholders on when to introduce L2 literacy acquisition, Nakamura and colleagues conducted studies across six language pairs in multiple regions in two countries with strikingly similar results: the presence of an L1 decoding threshold for L2 reading acquisition (Nakamura et al., 2019a, b). The first study was conducted in Karnataka and Andhra Pradesh states in South India, in Kannada-English and Telugu-English biliteracy acquisition pairs in grade 1–5 students. Utilizing a structural break analytical model precisely to determine if there is a point along the acquisition trajectory of L1 decoding skills where there is a “break” in the steepness of the slope of the relationship between L1 and L2 decoding, we found evidence for such a “threshold” point (60% on the L1 decoding test) in both states. In other words, we found a non-linear relationship between L1 and L2 decoding scores, in which there was a weak to non-existent correlation between Kannada and English scores below this threshold and a significant, strong, and exponentially increasing relationship above the threshold. This provided strong evidence for a non-linear transfer mechanism between L1 and L2 decoding skills in these alphasyllabic-alphabetic biliteracy acquisition pairs.

It is important to stress that the L1 Kannada and Telugu decoding tests were carefully constructed in a way that 60% of the test was precisely the point at which phonemic diacritic complexity increased in the syllabic structures of the akshara orthography, which in turn implies that when mastery of the phonemic components is attained, transfer readiness is triggered. This is consistent with the notion that phonemic awareness is the core shared resource between akshara (alphasyllabic) and English (alphabetic) reading development (Nakamura et al., 2014; Reddy & Koda, 2013).

A study was also conducted utilizing a similar methodology in Ethiopia (Nakamura et al., 2019a, b). This study was also motivated by stakeholders asking the same question for their educational contexts: when are children ready to benefit from an introduction to L2 reading instruction. This study approached the question in a similar way by investigating whether – and where – there may be a structural break in the linearity of the relationship between L1 and L2 decoding in four Ethiopian languages pairs: Afaan-Oromo and English; Amharic and English, Berta and English, and Wolayttatto and English. These pairs of languages were selected as they represent four different language families in Ethiopia (Zealelem, 2012). Amharic utilizes a Ge’ez script known as *fidel*. Like the akshara scripts in India (Nag, 2007; Nakamura et al., 2018), fidel also has an alphasyllabic structure in which the base unit is a syllable with phonemic graphemes represented within the syllable (Asfaha et al., 2009a, b). The other three languages – Afaan-Oromo, Berta, and Wolayttatto – all utilize the Roman alphabet. The results of this structural break analysis also revealed the presence of a threshold point of L1 decoding transfer readiness in all four language pairs. In line with the India study, the Amharic-English decoding relationship threshold break occurred at the point at which acquisition of the phonemic component within the fidel syllable block had been mastered, and in Afaan-Oromo, Berta, and Wolayttatto, the structural break was a little earlier in the L1 decoding acquisition process, reflecting the shared properties of the two

scripts. This also underscores the importance of L1-L2 distance in the non-linear transfer facilitation process.

Both these studies also tested for the presence of a structural break between L2 oral language proficiency and L2 decoding, but were unable to detect such non-linearity in the associations between these two sets of skills. Indeed, there were strong and significant relationships between L2 oral language skills and L2 decoding outcomes; however, the relationship was relatively linear. This implies “the more the better” for L2 oral language skills contributing to L2 reading, which was in contrast to the presence of a specific point of readiness for cross-language transfer of L1 decoding skills to L2 decoding. This is consistent with the constrained-unconstrained dichotomy, in which oral language skills are more likely to be unconstrained and thus perhaps less susceptible to a specific point at which mastery is considered “sufficient” for supporting L2 reading development.

These findings have been applied in the development of a full literacy teaching package in Laos and Guatemala in the form of formative assessments. While the biliteracy threshold component was only one part of an entire package, and thus we would not be able to pinpoint effectiveness of that component separately, mixed-method evaluations found significant positive impact on literacy development in both countries (Kamioka et al., 2018). In addition, there is ongoing research to determine if thresholds exist in a variety of reading sub-skills, including phonological awareness, decoding, and oral language comprehension sub-skills. These studies are underway in Côte d’Ivoire, Democratic Republic of Congo, Senegal, Mozambique, Philippines, and Kenya.

Together, these transfer facilitation threshold studies provide strong and consistent cross-country, cross-language pair evidence for a non-linear relationship between L1 and L2 decoding – especially in alphasyllabic-alphabetic and alphabetic-alphabetic biliteracy pairs. This evidence provides empirically driven recommendations to stakeholders to answer the question of “how much L1 is good enough” for L1-L2 decoding transfer, and is a promising approach to support decision makers in designing biliteracy programs in some LMIC contexts.

7 Conclusion

In this chapter, I set out to shine a light on the boundaries that exist between biliteracy theory, LoI policies and practices, and education stakeholder knowledge and demands in LMICs. These boundaries have been bottlenecks for the flow of information and knowledge that could help alleviate a dire situation of reading poverty in many educational communities across the world. By intentionally crossing boundaries and utilizing stakeholder input to inform empirical research, and vice versa, there is immense potential for bilingual reading theory to support language-in-education policy decision making – especially in terms of identifying the grades and skills required for successful transition from one language to another – across LMICs.

The current landscape on reading transfer theory is being expanded to account for issues such as what skills constitute transfer processes, how a variety of metalinguistic skills operate differently in transfer processes, how transfer might work bi-directionally in younger children, and what the interactive processes are in the complex mechanisms involved. The current landscape on LoI transitioning in LMICs remains focused on the core question of when to introduce L2 literacy instruction in bilingual and multilingual educational environments. These environments are mostly characterized by very limited classroom and teacher resources, but rich bilingualism and multilingualism, which could provide a strong candidate for available cognitive and linguistic resources to nurture biliteracy skills. These environments are also characterized by the socioeconomically- and politically- driven dilemma between the desire for more education in an L1 and persistent demand for earlier and earlier instruction in later acquired languages.

Within this space, it is worthwhile to consider how we can push the theoretical boundaries of transfer to conceptualize some corresponding sub-skills – such as decoding – as being non-linearly related. This methodological and theoretical expansion of the notion of transfer is borne out of a direct demand from stakeholders on timing the introduction of L2 literacy skills and thus is set up in the space between theory and policy in biliteracy acquisition. To be clear, this line of research is highly nascent. Future research is needed to examine other sub-skills that may be prime candidates for transfer thresholds, such as phonological or morphological awareness; to test transfer thresholds in a variety of language pairs; and to cross other geographical and contextual boundaries as well.

In sum, the need to apply cognitive transfer theory to macro educational policy and practice is especially acute in multilingual LMIC settings where teaching and classroom resources are limited and reading poverty is high. By focusing on the cognitive and linguistic resources children do have, it is possible harness available resources required to bridge the inequities related to language and literacy learning in LMICs.

References

- Abu-Rabia, S., & Siegel, L. S. (2002). Reading, syntactic, orthographic, and working memory skills of bilingual Arabic-English speaking Canadian children. *Journal of Psycholinguistic Research*, 31(6), 661–678.
- Alidou, H., Aliou, B., Brock-Utne, B., Diallo, Y. S., Heugh, K., & Wolff, E. (2006). *Optimizing learning and education in Africa – The language factor: A stock-taking research on mother tongue and bilingual education in sub-Saharan Africa*. UNESCO Institute for Education. Retrieved from: http://www.adeanet.org/adea/biennial-2006/doc/document/B3_1_MTBLE_en.pdf
- American Institutes for Research. (2022). *Strengthening bilingual and multilingual learning systems in Francophone Africa*.
- Annamalai, E. (2005). Nation-building in a globalised world: Language choice and education in India. In A. Lin & P. Martin (Eds.), *Decolonisation, globalization: Language-in-education policy and practice* (pp. 20–37). Multilingual Matters.

- Asfaha, Y. M., Beckman, D., Kurvers, J., & Kroon, S. (2009a). L2 reading in multilingual Eritrea: The influences of L1 reading and English proficiency. *Journal of Research in Reading*, 32(4), 351–365.
- Asfaha, Y. M., Kurvers, J., & Kroon, S. (2009b). Grain size in script and teaching: Literacy acquisition in Ge'ez and Latin. *Applied PsychoLinguistics*, 30(4), 709–724.
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Lawrence Erlbaum Associates.
- Azam, M., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change*, 61(2), 335–367.
- Beeharry, G. (2021). The pathway to progress on SDG 4 requires the global education architecture to focus on foundational learning and to hold ourselves accountable for achieving it. *International Journal of Educational Development*, 82, 102375.
- Bialystok, E. (2018). Bilingual education for young children: Review of the effects and consequences. *International Journal of Bilingual Education and Bilingualism*, 21(6), 666–679.
- Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A., & Sandefur, J. (2013). Scaling up what works: Experimental evidence on external validity in Kenyan education. *Center for Global Development Working Paper 321*.
- Branum-Martin, L., Mehta, P. D., Fletcher, J. M., Carlson, C. D., Ortiz, A., Carlo, M., & Francis, D. J. (2006). Bilingual phonological awareness: Multilevel construct validation among Spanish-speaking kindergarteners in transitional bilingual education classrooms. *Journal of Educational Psychology*, 98(1), 170–181.
- Brock-Utne, B. (2021). Language-in-education policies and practices in Africa with a special focus on Tanzania and South Africa. In J. Zajda (Ed.), *Third international handbook of globalisation, education and policy research* (pp. 609–627). Springer.
- Brown, J., Cartridge, R., Davies, S., Ul Hassan, E., Hsia, C., Lal, T., Mackison, D., Menon, R., Moratti, M., Murunga, V., Le Nestour, A., Whitehead, L., & White, H. (2018). *Research uptake in policymaking: From papers to policy CEDIL Inception Paper 14*.
- Burnett, N. (2019). Invited essay: It's past time to fix the broken international architecture for education. *International Journal of Educational Development*, 68(C), 15–19.
- Chakraborty, T., & Bakshi, S. K. (2016). English language premium: Evidence from a policy experiment in India. *Economics of Education Review*, 50, 1–16.
- Chung, S. C., Chen, X., & Geva, E. (2019). Deconstructing and reconstructing cross-language transfer in bilingual reading development: An interactive framework. *Journal of Neurolinguistics*, 50, 149–161.
- Coleman, H. (2011). Developing countries and the English language: Rhetoric, risks, roles and recommendations. In H. Coleman (Ed.), *Dreams and realities: Developing countries and the English language* (pp. 9–21). British Council.
- Collier, V. P., & Thomas, W. P. (2017). Validating the power of bilingual schooling: Thirty-two years of large-scale longitudinal research. *Annual Review of Applied Linguistics*, 37, 203–217.
- Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology*, 91(1), 29–43.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49(2), 222–251.
- Cummins, J. (1981). Empirical and theoretical underpinnings of bilingual education. *Journal of Education*, 163(1), 16–29.
- Da Fontoura, H. A., & Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese-English Canadian children. *Reading and Writing*, 7(1), 139–153.
- Deacon, S. H., Wade-Woolley, L., & Kirby, J. R. (2009). Flexibility in young second-language learners: Examining the language specificity of orthographic processing. *Journal of Research in Reading*, 32(2), 215–229.

- Donato, R., & Tucker, G. R. (2010). *A tale of two schools: Developing sustainable early foreign language programs*. Multilingual Matters.
- Edwards, S. (2019, December 10). *Rwanda's education U-turn sparks debate on mother tongue instruction*. Retrieved from: <https://www.devex.com/news/rwanda-s-education-u-turn-sparks-debate-on-mother-tongue-instruction-96129>
- Evans, D. K., & Mendez Acosta, A. (2021). Education in Africa: What are we learning? *Journal of African Economies*, 30(1), 13–54.
- Geva, E., & Siegel, L. S. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing*, 12(1), 1–30.
- Geva, E., Wade-Woolley, L., & Shany, M. (1997). Development of reading efficiency in first and second language. *Scientific Studies of Reading*, 1(2), 119–144.
- Gholamain, M., & Geva, E. (1999). Orthographic and cognitive factors in the concurrent development of basic Reading skills in English and Persian. *Language Learning*, 49(2), 183–217.
- Gottardo, A., Javier, C., Farnia, F., Mak, L., & Geva, E. (2014). Bidirectional cross-linguistic relations of first and second language skills in reading comprehension of Spanish-speaking English learners. *Written Language & Literacy*, 17(1), 62–88.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge University Press.
- Habitat for Humanity. (2022, February 13). *What is a slum? Definition of global housing crisis*. Retrieved from: <https://www.habitatforhumanity.org.uk/what-we-do/slum-rehabilitation/what-is-a-slum/>
- Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3), 607–668.
- Heugh, K. (2012). Cost implications of the provision of mother-tongue and strong bilingual models of education in Africa. In A. Ouane & C. Glanz (Eds.), *Optimising learning, education and publishing in Africa: The language factor* (pp. 255–289). UNESCO.
- Heugh, K., Benson, C., Bogale, B., & Yohannes, M. A. G. (2007). *Final report: Study on medium of instruction in primary schools in Ethiopia*. Ministry of Education.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2(2), 127–160.
- Hoover, W. A., & Tunmer, W. E. (2020). The cognitive foundations of reading acquisition. In W. A. Hoover & W. E. Tunmer (Eds.), *The cognitive foundations of reading and its acquisition: A framework with applications connecting teaching and learning* (pp. 41–84). Springer.
- Kamioka, N., Nakamura, P. R., & Stone, R. (2018). *Child literacy development pilot initiative and research in Laos and Guatemala*. American Institutes for Research and Catholic Relief Services. Retrieved from: <https://www.air.org/sites/default/files/Child-Literacy-Development-Pilot-Initiative-and-Research-Laos-Guatemala-2019.pdf>
- Karlsen, J., Lyster, S. A. H., & Lervåg, A. (2017). Vocabulary development in Norwegian L1 and L2 learners in the kindergarten–school transition. *Journal of Child Language*, 44(2), 402–426.
- Ke, S., Miller, R. T., Zhang, D., & Koda, K. (2021). Crosslinguistic sharing of morphological awareness in biliteracy development: A systematic review and meta-analysis of correlation coefficients. *Language Learning*, 71(1), 8–54.
- Kim, Y. S. (2009). Crosslinguistic influence on phonological awareness for Korean–English bilingual children. *Reading and Writing*, 22(7), 843–861.
- Kim, Y. S. G., & Piper, B. (2019). Cross-language transfer of reading skills: An empirical investigation of bidirectionality and the influence of instructional environments. *Reading and Writing*, 32(4), 839–871.
- Kim, Y. S. G., Lee, H., & Zuilkowski, S. S. (2020). Impact of literacy interventions on reading skills in low-and middle-income countries: A meta-analysis. *Child Development*, 91(2), 638–660.
- Koda, K. (2008). Impacts of prior literacy experience on second-language learning to read. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 68–96). Routledge.

- Koda, K., & Reddy, P. (2008). Cross-linguistic transfer in second language reading. *Language Teaching, 41*, 497–508.
- Kohler, M. (2017). *Review of language education policies internationally: Report commissioned by the multicultural education and languages committee (MELC)*. Government of South Africa.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.
- Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading, 34*(1), 114–135.
- Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first-and second-language learners. *Psychological Bulletin, 140*(2), 409–433.
- Ministry of Education, India. (2021). *National initiative for proficiency in reading with understanding and numeracy (NIPUN Bharat): A national mission on foundational literacy and numeracy*. Retrieved from: https://www.education.gov.in/sites/upload_files/mhrd/files/nipun_bharat_eng1.pdf
- Nag, S. (2007). Early reading in Kannada: The pace of acquisition of orthographic knowledge and phonemic awareness. *Journal of Research in Reading, 30*(1), 7–22.
- Nag, S., Chiat, S., Torgerson, C., & Snowling, M. J. (2014). *Literacy, foundation learning and assessment in developing countries*. Department for International Development.
- Nag, S., Vagh, S. B., Dulay, K. M., & Snowling, M. J. (2019). Home language, school language and children's literacy attainments: A systematic review of evidence from low-and middle-income countries. *Review of Education, 7*(1), 91–150.
- Nakamura, P. R., & de Hoop, T. (2014). *Facilitating reading acquisition in multilingual environments in India (FRAME-India). Final Report*. American Institutes for Research.
- Nakamura, P. R., Koda, K., & Joshi, R. M. (2014). Bilingual acquisition in Kannada and English: A developmental study. *Writing Systems Research, 6*(1), 132–147.
- Nakamura, P. R., Carson, K., Davis, D., & Rai, N. (2017). *Vamos Ler language mapping report*. American Institutes for Research.
- Nakamura, P. R., Joshi, R. M., & Ji, X. R. (2018). Investigating the asymmetrical roles of syllabic and phonemic awareness in akshara processing. *Journal of Learning Disabilities, 51*(5), 499–506.
- Nakamura, P. R., Bonilla, J., Mekonnen, B., Tefera, D., Gebrekidan, Z., Turner, K., Areaya, S., Elmeski, M., & Drummond, T. (2019a). *USAID reading for Ethiopia's achievement developed monitoring and evaluation (READ M&E): Language transition report*. American Institutes for Research.
- Nakamura, P. R., de Hoop, T., & Holla, C. U. (2019b). Language and the learning crisis: Evidence of transfer threshold mechanisms in multilingual reading in South India. *The Journal of Development Studies, 55*(11), 2287–2305.
- National Institute of Child Health, & Human Development (US) (NICHD), National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
- Olshtain, E., & Nissim-Amitai, F. (2004). Curriculum decision making in a multilingual context. *International Journal of Multilingualism, 1*(1), 53–64.
- Ouane, A., & Glanz, C. (2010). *Why and how Africa should invest in African languages and multilingual education: An evidence- and practice-based policy advocacy brief*. UNESCO Institute for Lifelong Learning.
- Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading, 7*(1), 3–24.
- Perfetti, C. A., & Dunlap, S. (2008). Learning to read: General principles and writing system variations. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 25–50). Routledge.

- Piper, B. (2017). Improving learning in sub-Saharan Africa using rigorous research designs. *Journal of Research on Educational Effectiveness*, 10(3), 535–540.
- Pretorius, E. J., & Mampuru, D. M. (2007). Playing football without a ball: Language, reading and academic performance in a high-poverty school. *Journal of Research in Reading*, 30(1), 38–58.
- Reddy, P. P. (2011). *Biliteracy development in middle-school students from the slums of India*. Unpublished doctoral dissertation. Carnegie Mellon University
- Reddy, P. P., & Koda, K. (2013). Orthographic constraints on phonological awareness in biliteracy development. *Writing Systems Research*, 5(1), 110–130.
- Rickard Liow, S. J., & Poon, K. K. (1998). Phonological awareness in multilingual Chinese children. *Applied PsychoLinguistics*, 19(3), 339–362.
- Scheele, A. F., Leseman, P. P., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied PsychoLinguistics*, 31(1), 117–140.
- Snow, C. E., & Matthews, T. J. (2016). Reading and language in the early grades. *The Future of Children*, 26(2), 57–74.
- Spolsky, B., & Shohamy, E. (2000). Language practice, language ideology, and language policy. In R. D. Lambert & E. Shohamy (Eds.), *Language policy and pedagogy: Essays in honor of a Ronald Walton* (pp. 1–43). John Benjamins Publishing Company.
- Tucker, G. R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (Eds.), *Beyond bilingualism: Multilingualism and multilingual education* (pp. 3–15). Multilingual Matters.
- United Nations. (2015). *Millennium development goals and beyond 2015*. Retrieved from: <https://www.un.org/millenniumgoals/>
- United Nations. (2022). *Sustainable development goals*. Retrieved from: <https://sdgs.un.org/goals>
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (1953). *The use of vernacular languages in education*. In Monographs on fundamental education VIII. UNESCO.
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (1990). *World declaration on education for all*. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000127583>
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (2000). *The Dakar framework for action: Education for all: Meeting our collective commitments (Including Six Regional Frameworks for Action)*. Retrieved from <https://unesdoc.unesco.org/ark:/48223/pf0000121147>
- United Nations Educational, Scientific and Cultural Organization (UNESCO) Institute for Statistics. (2017). *More than one half of children and adolescents are not learning worldwide*. Retrieved from: <http://uis.unesco.org/sites/default/files/documents/fs46-more-than-half-children-not-learning-en-2017.pdf>
- United States Agency for International Development (USAID). (2019). *USAID education: Early learning and foundational skills*. Retrieved from: https://www.usaid.gov/sites/default/files/documents/1865/USAID_Education_EarlyLearning_FactSheet_Apr19.pdf
- Verhoeven, L. T. (1994). Transfer in bilingual development: The linguistic interdependence hypothesis revisited. *Language Learning*, 44(3), 381–415.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing ability: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, 30, 73–87.
- Wang, M., Park, Y., & Lee, K. R. (2006). Korean-English biliteracy acquisition: Cross-language phonological and orthographic transfer. *Journal of Educational Psychology*, 98(1), 148–158.
- Wang, M., Yang, C., & Cheng, C. (2009). The contributions of phonology, orthography, and morphology in Chinese-English biliteracy acquisition. *Applied PsychoLinguistics*, 30(2), 291–314.
- Wawire, B. A., & Kim, Y. S. G. (2018). Cross-language transfer of phonological awareness and letter knowledge: Causal evidence and nature of transfer. *Scientific Studies of Reading*, 22(6), 443–461.

- World Bank. (2014). *Education for all*. Retrieved from: <https://www.worldbank.org/en/topic/education/brief/education-for-all>
- World Bank. (2017). *World development report 2018: Learning to realize education's promise*. The World Bank.
- World Bank. (2021a). *Loud and clear: Effective language of instruction policies for learning*. The World Bank.
- World Bank. (2021b). *Learning losses from Covid-19 could cost this generation of students close to \$17 trillion in lifetime earnings*. Retrieved from: <https://www.worldbank.org/en/news/press-release/2021/12/06/learning-losses-from-covid-19-could-cost-this-generation-of-students-close-to-17-trillion-in-lifetime-earnings>
- Zealelem, L. (2012). The Ethiopian language policy: A historical and typological overview. *Ethiopian Journal of Languages and Literatures*, 12(2), 1–59.