PRONUNCIATION AND ASSESSMENT

PROCEEDINGS OF THE 4$^{\text{TH}}$ ANNUAL PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING CONFERENCE

**Held at Simon Fraser University (Vancouver Campus)
Vancouver, British Columbia, Canada
August 24-25, 2012**

**Editors
John Levis
Kimberly LeVelle**

Table of Contents

Part V. Acquisition Studies

Part VI. Results of Instructional Treatments

Part VII. Teacher Development Practices

Levis, J., & LeVelle, K. (2013). Pronunciation and assessment. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference.* Aug. 2012. (pp. i-xii). Ames , IA: Iowa State University.

PRONUNCIATION AND ASSESSMENT

John Levis, Iowa State University
Kimberly LeVelle, Iowa State University

The 4th Pronunciation in Second Language Learning and Teaching conference went on the road for the first time, moving not only to another city, but to another country. The conference was held in Vancouver, British Columbia, Canada. Murray Munro of Simon Fraser University (SFU) was the host of the conference, which was held at SFU's Morris J. Wosk Centre for Dialogue in downtown Vancouver. All sessions were held in the beautiful building with a large conference room in the round, giving the feel of our own United Nations. The conference attracted over 100 participants from 18 countries, a significant increase in international participation over the three previous conferences.

The theme of the conference was Pronunciation in Language Assessment. Pronunciation is involved in the assessment of oral language proficiency and performance, but in scoring such tests, pronunciation is sometimes explicitly included by test creators and sometimes ignored. In addition, pronunciation ability must also be assessed in order to create an accurate picture of student needs. These varied uses of assessment are rarely addressed openly, but it was our hope that the conference would provide a venue for such discussion. The 4th annual conference included 34 concurrent sessions, 26 poster sessions (13 each day), a plenary address by Pavel Trofimovich of Concordia University in Montreal, and a Language Learning sponsored roundtable on Pronunciation Assessment. The conference schedule is below.

FRIDAY, August 24, 2012

08:15-08:50	Registration check-in (Atrium) and poster setup (Atrium/APH)	
08:50-09:00	Welcome (APH): Tom Perry, Chair, Department of Linguistics, Simon Fraser University	
09:00-10:10	**Plenary: Teaching second language pronunciation: From the psycholinguistic lab to the language classroom** *Pavel Trofimovich, Concordia University* (APH); Chair: Tracey Derwing	
10:10-10:35	*Break*	
Fri. AM	**FR1. Pedagogical Research (APH)**, R. Thomson, Chair	**FR2. Curriculum & Teaching (Rm 420)** G. Muller Levis, Chair
10:35-11:00	*Foote & McDonough* Using auditory priming tasks to target AWL word stress patterns	*Sardegna & McGregor* Scaffolding students' self-regulated efforts for effective pronunciation practice
11:05-11:30	*Saito* Recasts in instructed second language speech learning	*Reed* Operationalizing curriculum objectives: Integrating and assessing listening and speaking in an IEP
11:35-12:00	*Derwing, Foote & Munro* Teaching old dogs new tricks: L2 pronunciation instruction after 19 years of English experience	*Edelstein* L1 and L2 learners in the college public speaking course
Friday PM 12:15 - 2:00	**Working Lunch + Poster Session A: Student & Post-doc Posters (Atrium/APH)**	

Alsadoon	A01. Perception and production of English vowels
Aly Bailey & Brandl	A02. Pronunciation instruction in the beginning Spanish classroom: A perceptual study
Chang & Weng	A03. Late ESL learners' difficulties of producing lax and tense vowels in English
Chen	A04. Perception of English lexical stress by Chinese native speakers: A critical review
Gordon, Darcy & Ewert	A05. Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom
Lege & Tanner	A06. The effect of pause duration on comprehensibility
Ullakonoja, Van Moere, Huhta, Alderson, Haapakangas, & Nieminen	A07. L2 learners' oral reading fluency development during extensive reading intervention
Wang	A08. Prosodic acquisition: tone, stress and intonation
Zhuang	A09. You will speak like an American: ESL learner's pronunciation improvement
Friday PM 12:15 - 2:00	**Working Lunch + Poster Session A: Non-student Posters (Atrium/APH)**
De Meo, Pettorino, Vitale, Cutugno & Origlia	A10. Imitation/self-imitation in a computer-assisted prosody training for Chinese learners of L2 Italian
Sturm	A11. Liaison in L2 French: The effects of instruction
Ou	A12. Intelligibility and comprehensibility of English lexical stress and EIL phonological cores
Zetterholm	A13. Teaching Swedish as a foreign language

Fri PM	FR3. Technological Innovations (APH), Y. Wang, Chair	FR4. Teaching & Teacher Training (Rm 420), B. Zielinski, Chair
2:00-2:25	*Chun & Yu* Visualization of tone and intonation for teaching and learning Mandarin Chinese	*Thomson* Teachers' beliefs and practices in pronunciation teaching: Confidently right or confidently wrong?
2:30-2:55	*Richards* Optimizing the acquisition of AWL word stress patterns via a principled web-based flashcard pedagogy	*LeVelle & Levis* Learning to teach pronunciation: Attitudes, images, and identity
3:00-3:25	*Okamura* The comparison of L2 speakers' evaluation of and machine evaluation of Japanese learners' English	*Kirkova-Naskova, Tergujeff, Frost, Henderson, Kautzsch, Levey, Murphy & Waniek-Klimczak* The English pronunciation teaching in Europe survey: Teacher training and assessment
3:30-3:50	*Break*	

Fri PM	FR5. Studies of Learners (APH), K. Saito, Chair	FR6. Teaching Practices and Teacher Training (Cont.)
3:50 - 4:15	*Wilson & Horiguchi* How accurately people follow articulation instructions	*Tergujeff* English pronunciation teaching practices in Finland
4:20 - 4:45	*McCrocklin & Link* Accent and identity: Fear of sounding native?	*Costa Kurtz dos Santos* Pronunciation in the perspective of trainee teachers: an analysis of curricular training reports
4:50 - 5:15	*O'Brien* Successful L2 pronunciation?	*Poisson* Can a pronunciation approach based on research findings have an impact on learners' performance?

5:30 -	***Reception: Segal Centre (Rm 420-430) in the Harbour Centre across the street.***

SATURDAY, August 25, 2012

08:00-08:30	Registration check-in (Main Atrium) and poster setup (Atrium/APH)
08:30-10:45	**PSLLT Roundtable on Pronunciation Assessment (APH), sponsored by *Language Learning*** *Panelists: April Ginther, Luke Harding, Sara Kennedy, Rebecca Hincks, Murray Munro* *Discussants: Sarah Fleming, Beth Zielinski,* Moderator/Chair: John Levis.
10:45-11:15	*Break*

Sat. AM	**SA1. Assessment (APH),** O. Kang, Chair	**SA2. Teaching Pronunciation (Rm 420),** T. Harada, Chair
11:15-11:40	***Zielinski, Yates & Pryor*** Assessing pronunciation: How judgements of intelligibility relate to IELTS pronunciation scale scores	***Acton, Burri, Teaman & Baker*** Preliminaries to haptic-integrated pronunciation instruction
11:45-12:10	***Danforth & St. John*** The glossary project	***Rogan*** Pronunciation and task-based instruction
12:15-12:40	***Kang*** Pronunciation features distinguishing examinees	***Baker*** Integrating pronunciation into content-based ESL instruction

Saturday PM 1:45 to 3:00	**Poster Session B: Non-student Posters (Atrium/APH)**
Eustice	B01. Fledgling phonologists
Fotovatnia & Shahini	B02. Using *Pronunciation Power 2* to improve Iranian EFL learners' consonant production
Gonzalez Lopez & Counselman	B03. The acquisition of L2 pronunciation of Spanish by novice learners
Harada & Sato	B04. Effects of minimal exposure to English in early childhood on phonemic perception
Kondo	B05. Phonological memory on L2 pronunciation skills
Miller & Szymanski	B06. Improving oral proficiency with technology: A give and take
Muller Levis	B07. Lexical and grammatical features associated with contrastive focus
Munro, Derwing & Saito	B08. English L2 vowel acquisition over seven years
Ou, Yeh & Chuang	B09. Units of analysis, intelligibility evaluation and phonological cores of EIL
Rauber, Kluge, Rato, & Santos	B10. Designing audio, visual and audiovisual perceptual training tasks with TP application software
Rojczyk	B11. Spontaneous phonetic imitation of L2 vowels in a rapid shadowing task
Wallen	B12. Accent modification group: Using a mixed group format to address pronunciation concerns
Watts & Huensch	B13. Integrated speaking, listening, and pronunciation: Are textbooks leading the way?

Sat. PM	**SA3. Phonetic Issues (APH),** X. Wu, Chair	**SA4. L2 Speech Rating and Related Issues (Rm 420),** M. Reed, Chair

2:00-2:25		**Tanner, Landon & Porter** How does prosodic error frequency influence NNES' and NES' comprehensibility ratings?
2:30-2:55		**Fullana** Exploring listeners' response times in the assessment of L2 speech
3:00-3:25	**de Moras** The role of frequency in the acquisition of L2 pronunciation: the example of the French liaison	**Kawase, Hannah & Wang** Effects of visual speech information on native listener judgments of L2 speech intelligibility
3:30-3:50	**Break**	
Sat PM	**SA5. Phonetic Issues (Cont.),** X. Wu, Chair	**SA6. L2 Speech Rating and Related Issues (Cont.),** K. Saito, Chair
3:50-4:15	**George** The development of /θ/, a variable geographic phonetic feature, during a semester abroad: The role of explicit instruction	**Koffi** Confusion research as a complement to intelligibility research
4:20-4:45	**Sato** Effect of English-medium instruction on the production of VOT by Japanese learners of English	**Lima** Fundamental considerations in developing an intelligibility test for nonnative teaching assistants
4:50-5:15	**Silveira** Pronunciation instruction and syllabic-pattern discrimination	**Gomes** Understanding Brazilian way of speaking English, in theory and practice
5:15-	**CLOSING**	

Sessions

[Language Learning](#) Roundtable

The Roundtable on Pronunciation Assessment brought together five scholars presenting different perspectives of the assessment of pronunciation skills. Panelists were [April Ginther](#) (Purdue), [Murray Munro](#) (Simon Fraser), Sarita Kennedy (Concordia), [Rebecca Hincks](#) (KTH Royal Institute of Technology), and [Luke Harding](#) (University of Lancaster). Discussants were Sarah Fleming (IELTS) and Beth Zielinski (Macquarie). The moderator for the discussion was [John Levis](#) (Iowa State). The papers provided a historical overview of how pronunciation has been assessed (Munro), how intelligibility is assessed in relation to high-stakes spoken language tests and how intelligibility and nativeness are often conflated in such tests (Harding), ways to measure how pronunciation's components differentially contribute to holistic judgments of spoken language (Ginther), how pronunciation contributes to the assessment of interactive spoken language tests (Kennedy), and a discussion and overview of how computers are being used to automatically assess pronunciation and spoken proficiency, along with a critique of such attempts (Hincks). Because the papers given in the Roundtable are being used for an article being submitted to a journal, they do not appear in the proceedings. Instead, each presenter's abstract is reproduced below to give a fuller flavor of the talks given during the roundtable.

<div align="center">

April Ginther (Purdue University)
Measuring Characteristics of Spoken L2 English

</div>

The use of computers for the administration of speaking tests has considerably eased the burden associated with the capture and analysis of speech, and reliable assessment of spoken language can be established with

the use of human ratings in association with the use of holistic scales. However, fully validating holistic speaking scales requires explication and understanding of the differential contributions of the components of oral performance (e.g., pronunciation, fluency, accuracy, vocabulary) and the interaction among these components at different levels of holistic scales. Praat, a computer program with which you can analyze, synthesize, and manipulate speech, is a tool that allows close analysis of temporal and acoustic components of oral production. While temporal variables associated with fluency (e.g., speech rate, mean syllables per run) are relatively easy to capture, the selection and quantification of variables associated with pronunciation and prosody pose difficult but incredibly interesting challenges. This presentation will discuss findings from completed and ongoing studies that have used Praat to examine the temporal and acoustic properties of L1 Chinese speakers' performance at different levels of the Oral English Proficiency Test, a semi-direct test used to screen prospective international teaching assistants at Purdue University. These studies have examined speech rate, mean length of run, filled and silent pauses, pause placement, vowel quality, vowel length, consonant voicing, stress assignment, and pitch contour. Ongoing efforts are investigating ways to quantify the intelligibility of responses to read-aloud and free response items. Although the use of acoustic analysis software programs requires training and commitment, their use holds great promise not only for validating holistic speaking scales but also, and more importantly, explicating and understanding the characteristics and development of L2 speaking ability.

Luke Harding (Lancaster University)
Nativeness or Intelligibility: locating the construct in pronunciation scales

A significant challenge in assessing pronunciation – particularly in English language testing contexts – is the existence of what Levis (2005) calls 'two contradictory principles': the nativeness principle (that learners should achieve a native-like accent in the L2) and the intelligibility principle (that learners' pronunciation should be understandable to a broad range of listeners). The language of many current pronunciation assessment scales demonstrates a shift away from the nativeness principle, with statements in criteria focusing either on intelligibility (e.g., "easily understood"), or on more abstract pronunciation goals (e.g. "correct", "accurate"). However this shift presents several challenges for test developers in gauging the validity and reliability of their pronunciation criteria. First, it raises the question of whether raters still draw on perceptions of nativeness in their judgements of pronunciation, even when native norms are not invoked in scales. Secondly, it raises the question of whether raters from different language backgrounds interpret intelligibility, or abstract terms like "correct", in the same way. This paper will discuss these challenges, drawing on examples from pronunciation rating scales to illustrate key points.

Rebecca Hincks (Royal Institute of Technology (KTH))
Pronunciation Assessment Using Speech Technology

After decades of research, language technologies finally entered the mass market in the fall of 2011 with the release of the iPhone 4, whose main innovation was the introduction of Siri, the virtual, speech-directed personal assistant. As we become more comfortable with speech interfaces, we can expect growing trust in their use for pedagogical purposes. Language technologies are, relatively speaking, better at assessing pronunciation than at teaching it. Speech recognition (ASR) can identify deviant phonemes, without being able to easily provide a learner with information about what needs to be adjusted in terms of articulation. My contribution to the round table will report on the research challenges faced by engineers designing pronunciation assessment systems. Current issues include the development of technological alternatives to

ASR for assessment, and the relation between computer pronunciation error detection and human ratings of pronunciation. I will also reflect on how individual practitioners
in the field could beneficially take advantage of existing language technologies, using as an example my own work giving feedback on pitch variation to Chinese speakers of English.

Sarita Kennedy (Concordia University)
All Together Now: assessing pronunciation and communication in interaction

A significant shift is slowly taking place in the assessment of second language pronunciation. Test rubrics and assessors are no longer solely targeting speakers' use of native-like pronunciation; there is greater emphasis on assessing speakers' ability to communicate their message effectively; in addition, there is a growing trend for speaking assessments to include pair or group speaking tasks (Taylor, 2006). This means that assessing pronunciation and effective communication is no longer straightforward, raising several important questions: Should a rater's understanding of interactive speech always be the default measure of assessment if the rater is not also an interlocutor? Can we identify a common core of second language pronunciation (Jenkins, 2000) which is generally intelligible to any listener? Should pronunciation assessment also target speakers' use of effective strategies for managing problems in understanding, whomever the listener? These questions have implications for key aspects of second language pronunciation assessment, such as how raters and interlocutors are selected and how the target of evaluation is determined. In this presentation, these issues will be discussed with the aim of suggesting a research agenda for the future.

Murray Munro (Simon Fraser University)
Pronunciation Testing and Classroom Research: What should we test and why?

Language testing is carried out for many purposes, including diagnosis of learner difficulties, measurement of classroom learning, and assessment of L2 proficiency. In pronunciation research, 'testing' is also used as a means of evaluating the effectiveness of laboratory training procedures, classroom techniques, and CALL approaches. Implicit in all classroom-based research on pronunciation is the assumption that the dependent variable under study has some sort of relevance to teachers, students, and the students' interlocutors. In particular, when we wish to argue that a particular technique or piece of software 'works,' we assume that it has brought about changes in learners' skills and that we have evaluated those changes in a meaningful way. However, it is not always clear that these assumptions are valid. An examination of the limited set of classroom-based studies of pronunciation indicates that a diverse range of outcome measures have been employed, including perception scores, accent ratings from trained and untrained listeners, segmental and prosodic accuracy scores as assigned by the researchers themselves, and various types of intelligibility and comprehensibility measures. We have yet to see a clear convergence of opinion among researchers about which of these are genuinely useful in terms of their relevance to L2 communicative ability. Resolving this problem requires that classroom researchers work in concert with testing specialists to establish valid procedures for both research and testing purposes.

Plenary Address

The plenary was given by Pavel Trofimovich of Concordia University in Montreal. Entitled "From the psycholinguistic lab to the language classroom: What are some of the most efficient ways of helping non-

native speakers improve their ability to speak a second language?", the talk examined how an understanding of alignment in speech can be used to provide more effective teaching and more accurate assessment of pronunciation. The talk first described research in which L1 interlocutors repeat language used by another in order to build on what the other said. This repetition is part of the process of alignment, involving words, grammar, phonetic realization, accent and speech rate.

The main question addressed by Trofimovich was whether L2 speakers show the same kind of alignment shown by L1 speakers. Results of several studies provided evidence that successful L2-L2 interactions almost always used a lot of repetition, which was one way that L2 speakers used to promote mutual intelligibility. L2 learners were said to heavily rely on repetition in all types of learning including pronunciation learning. The talk then asked whether teachers could promote alignment in pronunciation through language tasks. A study using academic word stress patterns suggests that the kinds of tasks used in class can promote the natural process of alignment. This process was related to the psychological process of priming. Finally, the speaker suggested ways to promote alignment in teaching pronunciation, suggesting the types of activities promoting alignment and repetition have not been sufficiently exploited.

The proceedings version of the plenary is a shortened version of his talk. For a more complete discussion of the issues involved, see his forthcoming article in the journal *Language Teaching* (published by Cambridge University Press).

Proceedings Papers

The proceedings papers are a representative sample of the types of papers given at the conference. This year's proceedings include 25 papers (a new record) from six different general categories: Assessment, Phonetic Research, Technology, Acquisition Studies, Results of Instructional Interventions, and Teacher Development Practices.

Assessment

Three papers are included in the Assessment category. Okim Kang, in "Relative Impact of Pronunciation Features on Ratings of Non-Native Speakers' Oral Proficiency," examines how various errors in pronunciation predict ratings of oral proficiency on the Cambridge ESL General English Examination. Kang argues that there was a clear hierarchy of errors in relation to oral proficiency ratings. This short report does not give a more complete set of results, which will be included in a longer article to be published soon.

In *"Versant* and Advanced L2 Speakers' Ratings of Japanese Learners' Oral English," Akiko Okamura compares the automatic proficiency assessments of Japanese students provided by the Versant test with evaluations of the proficiency made by advanced proficiency L2 speakers in Sweden. Pronunciation seemed to play an important role in the judgments made by the Swedish L1 judges, but so did vocabulary.

The final paper in this category, "Teachers' Views on their Professional Training and Assessment Practices: Selected Results from the English Pronunciation Teaching in Europe Survey" is a contribution from a diverse group of pronunciation scholars and teachers from Europe (Anastazija Kirkova-Naskova, Elina Tergujeff, Dan Frost, Alice Henderson, Alexander Kautzsch, David Levey, Deirdre Murphy, and Ewa Waniek-Klimczak). The group looks at the assessment and teacher training findings from a larger survey. Teachers on the whole reported that they were not satisfied with the kind of training they received in

teaching pronunciation, and the ways that teachers assess pronunciation was not based on established scales. Indeed, it seemed somewhat ad hoc in how it was done.

Phonetics Research

Several paper took a phonetics-oriented approach to L2 pronunciation issues. In "Acoustic Cues for English Lexical Stress Perception by Mandarin Native Speakers: A Critical Review," Sibo Chen discusses Mandarin learners' perception of English lexical stress. The paper looks at previous studies of the clues used by Mandarin learners of English, identifies factors underlying disagreements in previous research, and suggests directions for pedagogy.

Ettien Koffi draws connections between L2 intelligibility research and another decades-old approach to listener understanding in "Confusion as a Complement to Intelligibility Research." Koffi describes the mechanics of confusion research before illustrating the ways confusion research operates by looking at three English vowels. He argues that confusion research, a model for making speech intelligibility prediction, is a valuable complement to current research into intelligibility.

Finally, Arkadiusz Rojczyk looks at how a rapid shadowing task affects the production of /æ/, non-categorical vowel in Polish that is prone to being confused with two other vowels, in *"Phonetic Imitation of L2 Vowels in a Rapid Shadowing Task."* The results showed that Polish learners of English modified their productions of the vowel due to exposure in the shadowing task, showing that L2 learners can successfully modify their production of unfamiliar vowels.

Technology

A regular feature of L2 pronunciation research is the use of technology to promote learning. Dorothy M. Chun, Yan Jiang, and Natalia Ávila examine the effectiveness of visual feedback in "Visualization of Tone for Learning Mandarin Chinese." Learners of Mandarin learned tones by comparing their production with the production of native speakers of Mandarin, both hearing the production and seeing the pitch tracks of the tone production. For the tones that were mispronounced in the pretest, nearly half improved in the posttest.

In "Imitation/Self-Imitation in a Computer-Assisted Prosody Training for Chinese Learners of L2 Italian," five Italian researchers (*Anna De Meo, Marilisa Vitale, Massimo Pettorino, Francesco Cutugno, and Antonio Origlia)* look at the use of "prosodic transplantation" to encourage more effective CALL training for pronunciation learning in Italian. Prosodic transplantation involves the use of native speaker prosodic patterns with the learners' own voices, and the researchers found that it yielded superior results to traditional imitation techniques.

In "Improving Oral Proficiency by Raising Metacognitive Awareness with Recordings," Jessica Miller looks at the use of voice recording technology to enhance metacognitive skills in German and French classes. The encouragement to help learners think more about their communication through self-evaluation had a positive impact on their Willingness-To-Communicate.

Acquisition Studies

A number of papers looked at L2 acquisition of different pronunciation features for different languages. In "English L2 Vowel Acquisition over Seven Years," Murray J. Munro, Tracey M. Derwing, and Kazuya Saito examined the rate of vowel learning for Mandarin and Slavic L1 learners of English over a seven-year period in Canada. They found that learning does not stop in naturalistic learning contexts, but that the rate of learning is greatest in the first year after arrival, after which development slows greatly.

Angela George examines social constraints on learning pronunciation features during study abroad in "The Development of /θ/, a Variable Geographic Phonetic Feature, During a Semester Abroad: The Role of Explicit Instruction." Although the feature being examined is one that exists in the learners' L1, George finds that their use of /θ/ actually decreased over time, which she attributes to attitudes toward the dialect being learned, dialect exposure and proficiency level.

In "Late ESL Learners' Difficulties of Distinction between Lax and Tense Vowels," Daniel Chang and Calvin Weng look at the accuracy of tense/lax vowel production by early- and late-bilingual Chinese learners of English. Mirroring other results in previous studies, they found that late-bilingual learners were more likely to mispronounce tense and lax vowels than were early-bilingual learners.

Discussion of prosodic difficulties in Swedish are discussed in "Prosodic Pitfalls when Learning Swedish as a Second Language" (Elisabeth Zetterholm). The paper uses recordings of learners of Swedish and discusses how long-time residents continue to have particular difficulties with vowel quantity and word stress.

In "Pronunciation Instruction and Syllabic-Pattern Discrimination," Rosane Silveira examines the vowel insertion of Brazilian learners of English, a syllabification strategy that adds extra syllables and potentially changes the word stress patterns of the users' speech. The study looked at the effect of instruction on learners' awareness of this strategy and their ability to distinguish between CVC and CV.CV words, and found that instruction regarding this common error was effective in improving perception.

Results of Instructional Treatments

A common theme at PSLLT is the way that instruction affects learning/acquisition. In "Liaison in L2 French: The Effects of Instruction," Jessica L. Sturm presents a study of whether phonetics training is effective in promoting greater use of liaison in French. Particularly interesting was the fact that instructed learners used fewer forbidden liaisons, a particular problem in acquiring liaisons.

In another study of liaisons, Nadine de Moras looks at the "The Role of Pronunciation Instruction on the Acquisition of *Liaisons* by Anglophone Speakers." Working from reports of near-native acquisition of liaisons by learners, de Moras compared the production of liaisons and enchainements by French native speakers and three groups of Anglophone speakers of French. All three Anglophone groups showed improvement by the posttest, but none reached a level that could be described as native-like.

"Scaffolding Students' Self-Regulated Efforts for Effective Pronunciation Practice" (Veronica G. Sardegna and Alison McGregor) examined how the use of practice and strategy use, along with teacher scaffolding, affected the accuracy scores of international teaching assistants in a 15-week course. Accuracy was measured for vowel reduction, linking, primary stress and intonation. The paper concludes that the role of teachers is critical for the success of self-regulated pronunciation practice.

"Pronunciation Teaching and Learning: Effects of Explicit Phonetic Instruction in the L2 Classroom," by Joshua Gordon, Isabelle Darcy and Doreen Ewert, tested how instruction in phonological features affected improvement in comprehensibility ratings. Results were interpreted in light of calls for attention to noticing and communicative approaches to teaching pronunciation.

Ann Aly Bailey and Anel Brandl look at the effect of instruction on early L2 learning in "Incorporating Pronunciation in the First-Year Spanish Classroom: An Early Intervention." Using a beginning L2 context, the authors look at the effects of technical and non-technical instruction as well as a control group receiving no instruction on pronunciation. The perceptual abilities of the non-technical instruction group showed significant differences at the posttest, but this difference faded at a delayed posttest.

Teacher Development Practices

The last category, Teacher development practices, is another consistent interest at PSLLT. Teachers are perennially reported to be under-trained to teach pronunciation, leading to interest in the causes and effects of this problem as well as how to more effectively train teachers. In "ESL Teachers' Beliefs and Practices in Pronunciation Teaching: Confidently Right or Confidently Wrong?" Ron Thomson examined the accuracy with which English language teachers evaluated online statements about pronunciation and how to teach and describe it. He found that teachers generally agreed with a core of uncontroversial statements, but that they were frequently uncertain about questionable or blatantly false claims.

William Acton, Amanda Baker, Michael Burri, and Brian Teaman discuss an unusual approach to pronunciation instruction and teacher training in "Preliminaries to Haptic-Integrated Pronunciation Instruction." Haptic techniques (movement plus touch) is presented as appropriate for use even by untrained instructors. It is also suggested that the techniques may be effective in promoting recall and integration of teaching targets when transferred to spontaneous speech.

In "Integrating Fluent Pronunciation Use into Content-Based ESL Instruction: Two Case Studies,"Amanda Baker looks in detail at how two teachers integrated pronunciation instruction into oral communication courses. Baker looks at five categories of pronunciation instruction: Language awareness, controlled practice, guided practice, fluency development and free practice. While all areas needed greater attention, fluency development was almost absent in teacher practice.

What do teachers actually think about pronunciation in the midst of teaching? Larissa Buss examines this question in "Pronunciation from the Perspective of Pre-Service EFL Teachers: An Analysis of Internship Reports." Her study of internship reports written by Brazilian EFL teachers suggested four broad themes: how teachers identify problems, how they explain them, how they address them, and finally, their beliefs about pronunciation and teaching. The results from these Brazilian teachers contrast with other results based on research looking at teachers in ESL contexts.

Patricia Watts and Amanda Huensch look at an understudied area of pronunciation teaching and teacher training in *"*Integrated Speaking, Listening and Pronunciation: Are Textbooks Leading the Way?" They looked at 11 integrated skills textbooks to see how pronunciation was integrated into oral

communication teaching. They suggest that, despite weaknesses, integrated textbooks do a relatively good job of reflecting the need for intelligibility-based goals in teaching pronunciation.

Finally, in "Understanding the Brazilian Way of Speaking English" Maria Lúcia de Castro Gomes describes the grass-roots formation of a professional study group among Brazilian teachers. The group studies phonetics, phonology and pronunciation teaching and explores the use of speech analysis technology to explore the acquisition of English pronunciation by Brazilian learners. The group offers a potential model for teachers who have no other access to consistent professional training in pronunciation.

Future Conference

When these proceedings come out, the 5th Pronunciation in Second Language Learning and Teaching Conference (again held at Iowa State University) will be history Please consider joining us next year at the University of California Santa Barbara, where the conference theme will be "Looking at L2 Pronunciation Research from Varying Perspectives." The call for submissions is at http://linguistlist.org/easyabs/PSLLT2014. **The final date for submissions is April 12, 2014.**

Sixth Annual Conference
Pronunciation in Second Language Learning and Teaching
University of California, Santa Barbara
Santa Barbara, California USA

Looking at L2 Pronunciation Research from Varying Perspectives
September 5-6, 2014
Plenary Speaker
Alene Moyer, University of Maryland

Pronunciation instruction is increasingly popular in language classrooms around the world, in second language and foreign language contexts. Issues of intelligibility (Munro & Derwing, 1995) vs. nativeness (Levis, 2005), functional load (Brown, 1991; Munro & Derwing, 2006), effective instructional techniques for overcoming learning plateaus (Acton, 1986; Hardison, 2004; Goodwin, 2006), fluency (Derwing et al., 2008) and the relative roles of suprasegmentals and segmentals in instruction (Hahn, 2004) have all been examined in multiple studies. However, a large majority of important research into pronunciation has been carried out with English as the target language, despite the importance L2 pronunciation in other languages, such as Japanese (e.g., Hirata, 2004), Spanish (e.g., Lord, 2008), French (Ruellot, 2006), German (Moyer, 1999), Chinese (Liu et al, 2000), and Dutch (Bongaerts, Mennen & Slik, 2000), among others. Research from a variety of L2 learning contexts is essential to filling out the current English-centric research agenda.

The 6th Pronunciation in Second Language Learning and Teaching Conference invites proposals for papers and posters on all topics related to naturalistic and classroom pronunciation acquisition and learning. We especially welcome proposals for papers on pronunciation in a wide variety of L2s other than English. Possible paper topics include descriptive and experimental studies, re-examinations of key research findings (e.g., intelligibility, comprehensibility and accentedness) in new languages, technology in the teaching of pronunciation, and innovative approaches to teacher education.

In addition to papers related to the place of pronunciation in L2s other than English, the conference invites proposals for papers or posters on any aspect of pronunciation research, teaching and learning. Papers will be given in English.

For further information about the conference, contact Dorothy Chun, Conference Organizer at pslltconference@gmail.com.

COMING IN 2015

THE JOURNAL OF SECOND LANGUAGE PRONUNCIATION

Published by John Benjamins Publishing Company

Editor in Chief: John Levis, Iowa State University (jlevis@iastate.edu)

The Journal of Second Language Pronunciation is a quarterly journal devoted to research into the acquisition, perception, production, teaching, assessment, and description of prosodic and segmental pronunciation related to second languages in all contexts of learning. *The Journal* publishes papers in four main areas:

- ·   experimental, instructed, and naturalistic research about second language pronunciation;
- ·   reviews and syntheses covering research perspectives on key pronunciation issues from different disciplines;
- ·   teaching-oriented perspectives on successful practices and research-based instruction;
- ·   reviews of technology and books focused on second language pronunciation.

The Journal encourages research that connects theory and practice, enhances our understanding of L2 phonological learning processes, and provides connections between L2 pronunciation and other areas of applied linguistics research such as pragmatics, CALL, and speech perception. Contributions focusing on empirical research will represent all portions of the methodological spectrum including quantitative, qualitative, and mixed methods studies.

The *Journal of Second Language Pronunciation* was born out of the professional goals and interactions of the PSLLT Conference and will be unique in providing a dedicated scholarly and interdisciplinary forum for research and practice into second language pronunciation. Research into second language pronunciation intersects with many other aspects of applied linguistics and with other fields, and such research is published in a wide variety of professional journals, professional proceedings, and other venues, yet no journal is yet dedicated to research on L2 pronunciation. The articles we envision as being appropriate for the journal include papers dealing with intelligibility and comprehensibility, accent, phonological acquisition, learning and teaching, the use of technology (such as automatic speech recognition, text-to-speech, and computer assisted pronunciation teaching), spoken language assessment, the social impact of L2 pronunciation, the ethics of pronunciation teaching, pronunciation acquisition in less commonly taught languages, language attitudes, speech perception and its relationship to speech production, and many other topics.

INTERACTIVE ALIGNMENT: IMPLICATIONS FOR THE TEACHING AND LEARNING OF SECOND LANGUAGE PRONUNCIATION

Pavel Trofimovich, Concordia University

Interactive alignment refers to a phenomenon characterized by interlocutors adopting and re-using each other's language patterns in the course of authentic interaction. According to the interactive alignment model, originally proposed by Pickering and Garrod (2004), this linguistic coordination in dialogue occurs at the level of the lexicon, grammar, and pronunciation and represents one way in which interlocutors achieve understanding in dialogue. The goal of this paper is to extend this model to second language (L2) pronunciation and to discuss its possible implications for L2 pronunciation teaching. Previous research suggests that repetition of language patterns at different linguistic levels is indeed a commonplace feature of communication among native speakers and can be observed and elicited in L2 learners. Therefore, pronunciation researchers and teachers might find interactive alignment as a useful framework to explain some of the complexities of L2 pronunciation development both from cognitive and sociocultural perspectives, as well as to develop or refine pedagogical activities for use in L2 pronunciation classrooms.

INTRODUCTION

The field of second language (L2) acquisition has seen increased interest in pronunciation research and its application to the teaching of pronunciation. Clearly, pronunciation can no longer be characterized as suffering from "the 'Cinderella syndrome'—kept behind doors and out of sight" (Celce-Murcia, Brinton & Goodwin, 1996, p. 323). A search in the *Linguistics and Language Behavior Abstracts* database for a period between 1996 and 2010, for example, yields 1,102 citations of studies investigating different aspects of L2 pronunciation, a nearly 40% increase from the number of studies (791) published on similar and related topics in the preceding 15-year period. At first glance, this sizeable body of knowledge does not compare favorably (at least in terms of quantity) with scientific output in other areas of L2 teaching and learning. For instance, the same database returns nearly three times the number of studies (2,787) published between 1996 and 2010 with a focus on L2 vocabulary. Nevertheless, the field of L2 pronunciation appears to be past its growing pains, with at least two regular international conferences (the one featured in this volume as well as the International Symposium on the Acquisition of Second Language Speech), numerous book-length volumes focusing on current theoretical thinking and pedagogical practices (e.g., Bohn & Munro, 2007; Hansen Edwards & Zampini, 2008; Reed & Levis, forthcoming; Levis & Munro, 2013), and thousands of active researchers and practitioners around the globe.

Another sign that the field L2 pronunciation has gained considerable ground comes from the number of theoretical proposals that researchers have advanced over the years to explain various aspects of L2 pronunciation. According to Thomas Kuhn, an influential American philosopher of science, this exemplifies what he termed normal science—a "puzzle-solving activity… a highly cumulative enterprise, eminently successful in its aim, [which is] the steady extension of the scope and precision of scientific knowledge" (1970, p. 52). Indeed, there is no shortage of interesting theoretical proposals designed to account for various aspects of L2 pronunciation

(e.g., Best & Tyler, 2007; Darcy et al., 2012; Eckman, 1991; Escudero & Boersma, 2004; Flege, 1995; 2007; Major, 2002; Trofimovich, Gatbonton & Segalowitz, 2007). However, what becomes apparent, especially to those involved in the practical task of language teaching, is that many of these proposals, while scientifically sound and engaging, often have little to contribute to L2 pedagogy. Fraser (2004) aptly captured this sentiment in her criticism of similar theoretical views:

> This is of course a valid scientific analysis, but is of limited direct use in the practical task of helping learners alter their pronunciation, since we are dealing here, not with phonological systems in collision, but with people learning a cognitive skill (p. 279).

While the utility of theories cannot (and should not) be judged solely based on their contributions to practice (see Teshakkori & Teddli, 2003), one goal of theory building in L2 pronunciation should be the establishment of "best practice", or the idea that research should ultimately inform (at least to a certain degree) L2 pronunciation teaching. This is because pronunciation is not simply a fascinating object of inquiry. Rather, pronunciation permeates all spheres of human life, lying at the core of oral language expression and embodying the way in which the speaker and the hearer work together to produce and understand each other's utterances. The goal of this paper is therefore to contribute to the overall objective of bridging the gap between research and practice by outlining a "teaching-friendly" view of L2 pronunciation learning and discussing its pedagogical applications. This view is an extension of the interactive alignment model to L2 pronunciation learning and teaching.

INTERACTIVE ALIGNMENT

Interactive alignment, as a theoretical view, originated in the field of cognitive psychology and was first articulated by Pickering and Garrod (2004). Underlying this view is the idea that dialogue is the most natural mode of human communication, and that the goal of interaction is for interlocutors to arrive at a common situation model. In other words, interlocutors need to establish "common ground", which includes (but is not limited to) information about people, time, actions, and their causes and consequences.

An interesting question here is how precisely interlocutors achieve such common ground in the course of an interaction. Pickering and Garrod proposed that at least one way of doing so is related to how interlocutors use language in the course of interaction. More specifically, interlocutors achieve understanding by aligning or coordinating their language at various levels: lexical, syntactic, and phonological. And this alignment becomes evident during conversation when interlocutors adopt and repeatedly use each other's language patterns. For example, native speakers engaged in communication tasks tend to re-use each other's lexical content and phrasal structure across turns as they work to construct a common understanding as part of interaction (e.g., Garrod & Anderson, 1987). This, Pickering and Garrod argued, illustrates convergence in language use which promotes successful communication.

Since then, researchers have shown that native-speaking interlocutors constantly demonstrate linguistic alignment or coordination in spoken interaction. Interlocutors re-use not only each other's words (Brennan & Clark, 1996; Garrod & Anderson, 1987) and grammatical structures (Branigan, Pickering & Cleland, 2000; Branigan, Pickering, Pearson & McLean, 2010) but also converge on common phonetic realizations of words (Clarke & Garrett, 2004; Pardo, 2006) and on common accent and speech rate (Giles, Coupland & Coupland, 1991). This re-use of language

patterns across interlocutors, indicative of alignment at various linguistic levels, has come to be seen as a powerful repetition-driven mechanism which supports successful interaction.

INTERACTIVE ALIGNMENT IN PRONUNCIATION

To date, the interactive alignment view has been successfully applied to describe different aspects of interaction between native speakers (Garrod & Pickering, 2009) and has been extended to bilingual code-switching (Kootstra, van Hell & Dijkstra, 2010). Can interactive alignment also be used to explain how learners acquire and use L2 pronunciation?

Alignment in native speaker interaction

There is now a considerable body of evidence that native-speaking interlocutors converge on common speech patterns in the course of interaction. This idea is far from being novel. In fact, what is referred to here as linguistic alignment at the level of pronunciation has been studied for decades within sociolinguistics as part of accommodation theory (Giles, 1973; Giles et al., 1991; Shepard, Giles & Le Poire, 2001). Briefly, accommodation theory is a framework for a study of linguistic and nonlinguistic behavior, in the context of social interaction, as a function of interlocutor beliefs, attitudes, and sociocultural conditions. Over 20 years ago, for example, Giles et al. (1991) listed several speech characteristics on which interlocutors appear to converge during laboratory-controlled and spontaneous interactions. These characteristics included utterance length, speech rate, information density, volume, pausing frequencies and lengths, as well as response latency. Accommodation theory explains such linguistic convergence as a sign of interlocutors' (often subconscious) desire for mutual social integration and identification and their need for mutual social approval.

More recently, working within the cognitive processing perspective, researchers have demonstrated tight links between interlocutors' speech output (production) and speech input (perception) in conversation. Pardo (2006), for instance, has shown that interlocutors converge on common phonetic realizations of words and that such convergence occurs rapidly (early on in the conversation) and persists for at least one week after the initial conversation. In another study, Kim, Horton and Bradlow (2011) have found that native-speaking interlocutors sharing the same dialect are more likely to converge on common phonetic and prosodic speech patterns than interlocutors with distinct dialects, suggesting that convergence is facilitated when interlocutors share a common linguistic background (see also Babel, 2010, 2012; Nielsen, 2011; Pardo, Jay & Krauss, 2010). Phonetic convergence can occur even for speech that is only seen. For example, listeners show similar degrees of phonetic convergence for words that they heard and for words that they lipread from a silent video recording of a speaker (Miller, Sanchez & Rosenblum, 2010). Taken together, these findings point to the existence of a rapid and likely automatic process of phonetic alignment in native-speaking interlocutors. This process appears to reflect a human perceptual system which adapts readily in response to recent experience (Samuel & Kralijc, 2009).

Alignment in L2 speakers

When it comes to L2 learners interacting with other learners or with native speakers, it is far less obvious whether and under what circumstances learners demonstrate interactive alignment in pronunciation. It appears, though, that phonetic convergence depends on several related variables, including the native language background of interlocutors and their familiarity with each other's way of speaking (Costa, Pickering & Sorace, 2008). For example, compared to two

native speaker interlocutors, a native speaker and an L2 learner in conversation show reduced phonetic convergence (Kim et al., 2010). Dialogue partners with a common language background are also those who demonstrate more successful interactions than linguistically mismatched interlocutors, as judged through the quality of information exchanged and task completion times (Van Engen et al., 2010). Thus, whether or not L2 learners align with their interlocutors seems to depend on learners' familiarity with their interlocutors' language background, such that a shared language background is associated with more phonetic alignment.

There is also some preliminary evidence that the extent to which L2 learners align with their interlocutors might be related to the degree of accent in learner speech. For instance, compared to learners with either strong or weak accents, only moderately-accented learners appear to show phonetic convergence with a native-speaking interlocutor (Kim et al., 2010). Assuming that accent ratings capture some aspects of L2 speaking proficiency, phonetic convergence may depend on learners' mastery of the L2 phonetic system and their perception of the interlocutor's communicative needs. In other words, learners whose accent is very non-nativelike may not have the linguistic means to align with their interlocutor, while learners with very nativelike accents may not perceive the need to align because communication is not compromised.

Further evidence for interactive alignment in pronunciation comes from two recent classroom-based studies with university-level L2 learners of English (Trofimovich, McDonough & Neumann, 2013; Foote, Trofimovich & McDonough, submitted). These researchers examined the effectiveness of communicative activities providing learners with L2 models (fully-formed, targetlike utterances) at two linguistic levels: grammar (passives) and pronunciation (word stress in multisyllabic words). According to the interactive alignment view (Pickering & Garrod, 2004), the degree of alignment—defined as learners' ability to re-use the provided model structures in completion of novel, self-generated utterances—should be enhanced when learners are repeatedly exposed to models with integrated patterns of grammar and pronunciation rather than simply models of grammar or of pronunciation. This is precisely what was observed. Only integrated models (word stress combined with the passive) were successful at eliciting alignment, or repetition of grammatical structure and stress (Trofimovich et al., 2013). In addition, communicative activities of this kind were effective at providing learners with opportunities for practice. Each learner generated, on average, 46 multisyllabic words as part of four brief communicative activities, and heard a similar number of words spoken by his or her interlocutor (Foote et al., submitted). Thus, alignment can be elicited through collaborative, communicative tasks which have potential for providing sustained practice of target pronunciation patterns.

IMPLICATIONS FOR PRONUNCIATION LEARNING AND TEACHING

If repetition of language patterns at different linguistic levels is indeed a commonplace feature of communication among native speakers and can be observed and elicited in L2 learners, then what can interactive alignment, as a theoretical view, offer to L2 pronunciation teaching? The answer to this question likely depends on a clear understanding of what underlies phonetic alignment in dialogue. In their original model, Pickering and Garrod (2004) proposed priming as the main mechanism of alignment in dialogue. Priming is essentially an implicit, unconscious repetition phenomenon. It refers to speakers re-using language patterns experienced in recent discourse (McDonough & Trofimovich, 2008). There is strong support for repetition and priming as implicit phenomena in the fields of social and cognitive psychology, both for native speakers and L2 learners. In social psychology, for instance, mimicry (i.e., verbal, facial, emotional, and

behavioural repetition) has been long regarded as an automatic and implicit phenomenon of social behavior (Chartrand & Dalton, 2008). And in the field of cognitive psychology, the unconscious repetition of language patterns experienced in recent discourse (shown as priming effects) is considered an automatic and implicit language learning mechanism (Ferreira & Bock, 2006; McDonough & Trofimovich, 2008). Thus, the involvement of implicit learning in linguistic alignment is established. What needs to be clarified, though, is how more explicit and overt ways of language learning and use relate to alignment and how such explicit ways of learning (e.g., category formation, inferencing) may be harnessed to promote linguistic alignment.

Although it may be premature to suggest definitive applications of the interactive alignment view to L2 pronunciation teaching, at least until we better understand the implicit and explicit mechanisms underlying repetition in dialogue, several possibilities nevertheless come to mind. First, L2 learners will likely benefit from awareness-raising activities that will sensitize them to the fact that successful interaction often involves a lot of repetition. Learners might also benefit from listening activities featuring authentic spoken interaction, in order to become aware of pronunciation patterns (both segmental and especially suprasegmental) often repeated between interlocutors. Learners may then become more sensitive to how repetition can be used to construct successful interactions (Bremer & Simonot, 1996; Watterson, 2008). Second, if we adopt the alignment view, then pronunciation activities specifically targeting linguistic alignment hold some promise in pronunciation teaching. This includes collaborative classroom-based activities designed to elicit alignment, activities featuring corrective feedback and especially recasts as repeated models of targetlike language patterns, as well as activities built around high-frequency, functional, formulaic language (e.g., Foote et al., submitted; Gatbonton & Segalowitz, 2005; Saito & Lyster, 2012; Trofimovich & Gatbonton, 2006; Trofimovich et al., 2013).

Third, if we assume that alignment is enhanced when learners encounter patterns of language that match in many ways—for example, in terms of grammar and pronunciation—we can also hypothesize that alignment should also be enhanced for patterns of language experienced simultaneously across several modalities, sensory channels, and presentation media. And there is some very interesting evidence emerging about the effectiveness of multimodal, multisensory techniques applied to the teaching of L2 pronunciation (e.g., Hardison, 2010; Levis & Pickering, 2004; Sueyoshi & Hardison, 2005). Last but not least, the alignment view implies that different kinds of imitation activities—such as silent mouthing (Davis & Rinvolucri, 1990), mirroring, echoing, shadowing (Celce-Murcia et al., 2010), as well as dramatic imitation techniques that involve imitating not only speech, but also gestures, facial expressions, and affect (Hardison & Sonchaeng, 2005)—may be particularly useful in helping L2 learners align to a model.

CONCLUSION

In his introduction to a 2005 special issue of *TESOL Quarterly* devoted to pronunciation, Levis wrote about changing contexts and shifting paradigms in L2 pronunciation research. He emphasized the crucial role of intelligibility, rather than nativeness, as a goal of pronunciation teaching, highlighted both the speaker and the hearer as being essential to communication, and underscored the important role of identity and by extension, of wider sociocultural context, in L2 interaction. Interactive alignment—as a theoretical framework applied to the teaching and learning of L2 pronunciation—has the potential to contribute to the changing paradigms and shifting contexts defined this way. First, the alignment view firmly places intelligibility as central to communicative success (Derwing & Munro, 2009; Levis, 2005). If interlocutors' goal

is to achieve understanding, then intelligibility problems can be viewed as failure to align at the level of phonetic/prosodic perception and production. Interactive alignment thus becomes one means for interlocutors to resolve and avoid communication breakdowns, particularly when intelligibility compromises smooth and efficient communication. Second, the alignment view firmly establishes pronunciation within communicative approaches to language learning and teaching (Celce-Murcia et al., 2010), with a dual focus on both the speaker and the hearer as active participants in communication. Finally, the alignment view does not exclude social and contextual influences on learning. For example, in the course of interaction, interlocutors might align not only in terms of language but also in terms of gestures, facial expressions, eye gaze, and body movement (Atkinson, Churchill, Nishino & Okada, 2007; Churchill, Nishino, Okada & Atkinson, 2010). In fact, alignment can be viewed even more broadly—in the context of an individual's interaction with his or her environment (Atkinson, 2011). It is possible to imagine, then, that interlocutors can also align (or fail to do so) at the level of social factors, such as attitudes, beliefs and identity, and that these could influence the nature of interaction and the quality of language produced (Lindemann, 2002). From this vantage point, the interactive alignment view emerges as a useful framework for researchers to explain some of the complexities of L2 pronunciation development both from cognitive and sociocultural perspectives, and for teachers to develop or refine activities for use in L2 pronunciation classrooms.

ACKNOWLEDGEMENTS

ABOUT THE AUTHOR

Pavel Trofimovich (PhD 2003, University of Illinois at Urbana-Champaign) is an Associate Professor of Applied Linguistics in the Department of Education at Concordia University in Montreal. He conducts research on the acquisition of second language pronunciation, speech perception/production, phonological processing, and the role of cognitive and social variables in second language learning. He has co-authored two books with Kim McDonough on the use of cognitive psycholinguistic research methods in second language research—*Insights from psycholinguistics: Applying priming research to L2 learning and teaching* (2011, John Benjamins) and *Using priming methods in second language research* (2008, Routledge). He has received funding from SSHRC, FQRSC, and the British Council (UK) and has published numerous peer-reviewed articles and book chapters. He currently serves as the director of the graduate program in Applied Linguistics at Concordia, an associate editor of *Language Learning*, and an editorial board member of *Language Learning and Technology.*

REFERENCES

Atkinson, D. (2011). A sociocognitive approach to second language acquisition: How mind, body, and world work together in learning additional languages. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 143-166). New York: Routledge.

Atkinson, D., Churchill, E., Nishino, T., & Okada, H. (2007). Alignment and interaction in a sociocognitive approach to second language acquisition. *Modern Language Journal*, *91*, 169-188.

Babel, M. (2010). Dialect convergence and divergence in New Zealand English. *Language in Society*, *39*, 437- 456.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, *40*, 177-189.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 13-34). Amsterdam: John Benjamins.

Bohn, O.-S., & Munro, M. J. (Eds.) (2007). *Language experience in second language speech learning: In honor of James Emil Flege.* Amsterdam: John Benjamins.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. *Cognition*, *75*, B13-25.

Branigan, H. P., Pickering, M. J., Pearson, J., & McLean, J. F. (2010). Linguistic alignment between people and computers. *Journal of Pragmatics*, *42*, 2355-2368.

Bremer, K., & Simonot, M. (1996). Joint negotiation of understanding: Procedures for managing problems of understanding. In K. Bremer, C. Roberts, M. Vasseur, M. Simonot, & P. Broeder (Eds.), *Achieving understanding: Discourse in intercultural encounters* (pp. 181-206). London: Longman.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482-1493.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J.M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages.* New York: Cambridge University Press.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J.M. (2010). *Teaching pronunciation: A course book and reference guide.* New York: Cambridge University Press.

Chartrand , T. L., & Dalton, A. (2008). Mimicry: Its ubiquity, importance, and functionality. In E. Morsella, J. Bargh, & P. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 458-483). Oxford: Oxford University Press.

Churchill, E., Nishino, T., Okada, H., & Atkinson, D. (2010). Simbiotic gesture and the sociocognitive visibility of grammar. *Modern Language Journal*, *94*, 234-253.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, 116, 3647-3658.

Costa, A., Pickering, M. J., & Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes*, *23*, 528-556.

Darcy, I., Dekydtspotter, L., Sprouse, R., Glover, J., Kaden, C., McGuire, M., & Scott, J. (2012). Direct mapping of acoustics to phonology: On the lexical encoding of front round vowels in L1 English-L2 French acquisition. *Second Language Research*, *28*, 5-40.

Davis, P., & Rinvolucri, M. (1990). *The confidence book: Building trust in the language classroom*. Essex: Longman.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, *42*, 476-490.

Eckman, F. R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, *13*, 23-41.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551-585.

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes*, *21*, 1011-1029.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 229-273). Timonium, MD: York Press.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole & J. Hualde (Eds.), *Laboratory Phonology 9* (pp. 353-380). Berlin: Mouton de Gruyter.

Foote, J. A., Trofimovich, P., & McDonough, K. (submitted). Interactive alignment during collaborative tasks and the production of AWL word stress patterns.

Fraser, H. (2004). Constraining abstractness: Phonological representation in the light of color terms. *Cognitive Linguistics*, *15*, 239-288.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*, 181-218.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*, 292-304.

Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*, *61*, 325-353.

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of accommodation: Developments in applied sociolinguistics* (pp. 1-68). Cambridge: Cambridge University Press.

Giles. H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, *15,* 87-105.

Hansen Edwards, J. G., & Zampini, M. L. (Eds.) (2008). *Phonology and second language acquisition.* Amsterdam: John Benjamins.

Hardison, D. M. (2010). Visual and auditory input in second-language speech processing [Research Timeline]. *Language Teaching*, *43*, 84-95.

Hardison, D. M., & Sonchaeng, C. (2005). Theatre voice training and technology in teaching oral skills : Integrating the components of a speech event. *System*, *33*, 593-608.

Kim, M., Horton, W. S., & Bradlow, A. R. (2010). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, *2*, 125-156.

Kootstra, G. J., van Hell, J. G., & Dijkstra, T. (2010). Syntactic alignment and shared word order in code-switched sentence production: Evidence from bilingual monologue and dialogue. *Journal of Memory and Language, 63,* 210-231.

Kuhn, T. S. (1970). *The structure of scientific revolutions*. Cihcago: The University of Chicago Press.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*, 369-377.

Levis, J., & Munro, M. (2013). Phonetics and phonology: Overview. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-5). Oxford: Wiley-Blackwell.

Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System, 32,* 505-524.

Lindemann, S. (2002). Listening with an attitude: A model of native-speaker comprehension of non-native speakers in the United States. *Language in Society, 31*, 419-441.

Major, R. C. (2002). The phonology of the L2 user. In. V. Cook (Ed.), *Portraits of the L2 user* (pp. 67-92). Clevedon: Multilingual Matters.

McDonough, K., & Trofimovich, P. (2008). *Using priming methods in second language research*. New York: Routledge.

Miller, R. M., Sanchez, K., & Rosenblum, L. D. (2010). Alignment to visual speech information. *Attention, Perception, and Psychophysics*, *72*, 1614-1625.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*, 132-142.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, *119*, 2382-2393.

Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, and Psychophysics, 72*, 2254-2264.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169-226.

Reed, M., & Levis, J. (Eds.). (forthcoming). *The handbook of English pronunciation*. London: Wiley-Blackwell.

Saito, K., & Lyster, R. (2012). Investigating the pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly*, *46*, 387-398.

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, Psychophysics*, *71*, 1207-1218.

Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication accommodation theory. In W. P. Robinson & H. Giles (Eds.), *The new handbook of language and social psychology* (pp. 33-56). New York: Wiley.

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*, 661-699.

Teshakkori, A., & Teddli, C. (Eds.) (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in L2 Spanish word processing: Implications for pronunciation instruction. *Modern Language Journal*, *90*, 519-535.

Trofimovich, P., Gatbonton, E., & Segalowitz, N. (2007). A dynamic look at L2 phonological learning: Seeking psycholinguistic explanations for implicational phenomena. *Studies in Second Language Acquisition*, *29*, 407-448.

Trofimovich, P., McDonough, K., & Neumann, H. (2013). Using collaborative tasks to elicit auditory and structural priming. *TESOL Quarterly*, *47*, 177-186.

Van Engen, K. J., Baese-Berk, M., Baker, R. E., Kim, M., & Bradlow, A. R. (2010). The Wildcat Corpus of native- and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles. *Language & Speech*, *53*, 510-540.

Watterson, M. (2008). Repair of non-understanding in English in international communication. *World Englishes*, *27*, 378-406.

Kang, O. (2013). Relative impact of pronunciation features on ratings of non-native speakers' oral proficiency. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference.* Aug. 2012. (pp. 10-15). Ames , IA: Iowa State University.

RELATIVE IMPACT OF PRONUNCIATION FEATURES ON RATINGS OF NON-NATIVE SPEAKERS' ORAL PROFICIENCY

Okim Kang, Northern Arizona University

In the context of second language (L2) assessment, pronunciation has proved to be a problematic and consequently not widely investigated area. Although some effects of fluency, suprasegmental, or individual segmental errors have been identified in L2 oral assessment, the relative impact of a wide array of pronunciation features on non-native speakers' (NNSs) oral proficiency is still uncertain. The current study has investigated to what extent errors in various pronunciation features predict the rated oral proficiency of NNSs' speech. Speech samples of 120 candidates, 1-2 minute long, taken from Individual Long Turn sections in Cambridge ESOL General English Examinations, were analyzed acoustically for measures of speech rate, pauses, stress, intonation, and segmental errors. Results showed that there was a hierarchical priority in the effect of the pronunciation features on ratings of NNSs' oral proficiency. The finding has important implications for the field of L2 speaking assessment and pronunciation instruction.

THE STUDY

Various aspects of non-native speakers' (NNSs) pronunciation can be considered in listeners' assessments of speaker proficiency and both segmentals and suprasegmentals have been examined for their roles in judgments of accented speech. Some studies have investigated the impact of suprasegmental features on listeners' ratings of NNSs' oral performance (Kang, Rubin, & Pickering, 2010) or the correlations between temporal measures and listeners' rating scores (Munro & Derwing, 2001). Other studies have examined the effect of segmentals on listeners' judgments of NNSs' speech (e.g., segmental errors show great effects on listeners' comprehension in Fayer & Krasinski, 1987; high functional-load segmental errors have larger effects on perceptual scales than low functional-load segmental errors in Munro & Derwing, 2006).

However, no consensus has been reached regarding to what extent different pronunciation features contribute to the overall ratings of speaking assessment. Using candidates' spoken responses on the Cambridge ESOL General English Examinations; this study examined the relative contribution of each of the pronunciation categories to the overall oral proficiency judgments.

Speech samples

Cambridge ESOL General English Examinations provided one hundred twenty speech files; these were candidates' responses to their prompts in speaking test. From each speech file, a one-minute-long, monologic section of the response was extracted and transcribed. The current study included the following four proficiency levels using the Common European Framework of Reference for Languages (CEFR) from B1-C2: the

Preliminary English Test (PET, B1), the First Certificate in English (FCE, B2), the Certificate in Advanced English (CAE, C1), the Certificate of Proficiency in English (CPE, C2). Speakers analyzed in this study included 32 in PET, 32 in FCE, 34 in CAE, and 22 in CPE. There were 21 different first languages (L1s) represented.

Speech data coding and analysis

This study auditorily and instrumentally analyzed segmental and suprasegmental features of candidates' output. The instrumental analysis was conducted on two versions of speech analysis software: the Computerized Speech Laboratory (CSL) for stress and intonation analysis and PRAAT for fluency analysis. For the auditory analysis, two analysts participated with reliability reaching .81.

Segmental (vowels and consonants) errors included vowel/consonant substitutions and fourteen other types of errors (e.g., simplification of consonant clusters, linking errors, vowel or consonant epenthesis, vowel or consonant elision). The segmental errors were calculated as the total number of segmental errors divided by the total number of syllables articulated.

The suprasegmental analysis employed the methodology from Kang et al. (2010), i.e., it consisted of fluency, stress and pitch, and tone choices. In fluency measures, the syllables per second, mean length of run (e.g., utterances between pauses of 0.1 seconds and above), the number of silent pauses, mean length of silent pauses, number of hesitation markers, and mean length of hesitation pauses were counted. To measure stress and pitch, space (the proportion of prominent words to the total number of words), pace (number of prominent syllables per run), primary stress (number of misplaced lexical stresses), and overall pitch range were examined.

In addition, nine tone choices were identified following Brazil's (1997) framework for the measures of pitch movement and level: high-rising, high-level, high-falling, mid-rising, mid-level, mid-falling, low-rising, low-level, and low-falling. In other words, the pitch measures included pitch levels of prominent syllables (high, medium, or low) and pitch movement within tone units (rising, level, or falling). Previous research has demonstrated the importance of tone choices on the focus word and pitch level as they can affect perceived information structure in L2 discourse (Kang et al., 2010; Pickering, 2001; Wennerstrom, 1994).

RESULTS

A step-wise multiple regression was performed with the four proficiency levels as a dependent variable and each of the four categories of pronunciation features as independent variables (segmental errors, fluency, stress and pitch, and tone choices). Table 1 illustrates the proportion of variance (R^2) for each of the four pronunciation categories.

Table 1

Proportion of Variance (R^2) Explained by Pronunciation Features

Pronunciation Features	R Square (R^2) Change
Fluency	.267
Stress and pitch	.309
Tone choice	.045
Segmental errors (high FL vowels and consonants, Low FL vowels and consonants, and 13 other segmentals)	.080
Total	.701

The results showed that 70% of the variance in Cambridge ESOL four proficiency levels was attributed to pronunciation errors. Among those features, fluency took up 26.7%, stress and pitch, 30.9%, tone choice, 4.5%, and segmental errors, 8%. Consequently, from the linguistic analysis perspectives in speaking assessment, we can hypothesize that there is a hierarchical priority in the effect of these pronunciation features (Figure 1) on L2 oral proficiency. That is, in assessing NNSs' oral performance, various pronunciation criterion features contribute to raters' judgments differently. Stress and pitch played a more important role than any other features, followed by fluency, segmental errors, and intonation tone choices.
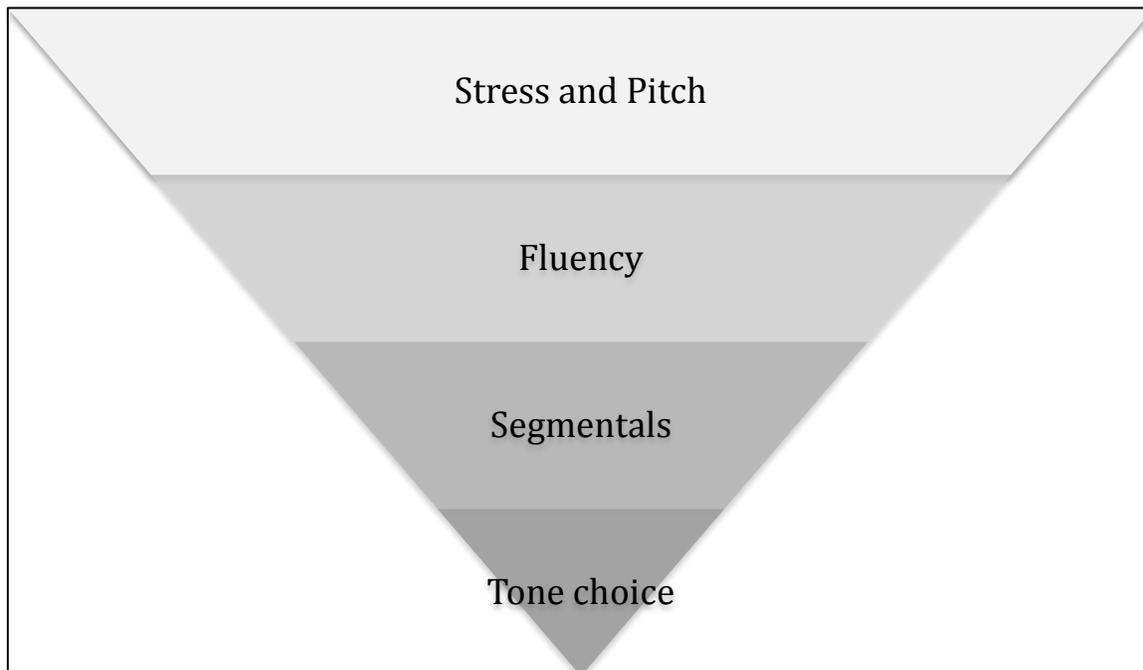
Figure 1. Hierarchical Priority in Pronunciation Features

DISCUSSION AND CONCLUSION

The results show the contribution of each of the pronunciation criterion features to the oral proficiency assessment. Note that most previous research has been limited in terms of the scope of the pronunciation features examined (e.g., Kang et al., 2010). In this study,

a variety of pronunciation variables (segmental errors, fluency, stress and pitch, and tone choices) were examined together and explained 70% of the variance. Fluency and stress/pitch features were strong contributors to this judgment (combined 56%) whereas segmental errors and intonation contributed more weakly to proficiency judgments.

One of the interesting findings is the strong contribution of stress and pitch variables on oral proficiency assessment. Thirty-one percent of the variance in Cambridge ESOL proficiency ratings was attributable to these stress and pitch factors, which is somewhat higher than the fluency dimension (27%). The significance of stress and pitch in NNSs' speech evaluation is in line with recent research examining pronunciation features on listeners' judgments of accented speech (Kang, 2010; Kang et al., 2010; Trofimovich & Isaacs, 2012). Trofimovich and Isaacs (2012) demonstrated that word stress influenced listeners' judgments (both accent and comprehensibility) in the case of rating the oral production of native French speakers of English. In Kang et al. (2010), stress was part of the acoustic fluency cluster, which was the most potent predictor of listeners' evaluation of NNSs' speech. In addition, pitch range was one of the strongest predictors of undergraduate students' ratings of an international teaching assistant's in-class lecture (Kang, 2010).

The importance of stress and pitch is particularly noteworthy. In speaking assessment, most studies have investigated the effects of temporal/fluency measures (speech rate and pauses) on listeners' judgments of NNSs' accented speech (e.g., Iwashita et al., 2008; Kormos & Denes, 2004; Ginther, Dimova, & Yang, 2010). These temporal measures are also frequently preferred by automatic speech recognition systems as objectively measurable parameters that show a high correlation with L2 fluency judgments (Zechner, Higgins, Xi, & Williamson, 2009). In this study, 27% of the variance in the proficiency ratings was explained by these fluency features. Nevertheless, the finding of the present study suggests that stress and pitch should be weighted more heavily in L2 pronunciation assessment. As Trofimovich and Isaacs (2012) stated, the influential contribution of stress to the oral judgments of accented speech can be related to "the fact that stress is one of the most structural and hierarchical aspects of phonology" (p. 10).

Overall, there appears to be a clear hierarchical structure in the importance of pronunciation features. In this study, stress and pitch were ranked first, followed by fluency measures, segmental errors, and tone choices at the end. Accordingly, ESL teachers may need to prioritize pronunciation features in classroom instruction for the promotion of intelligibility or for the preparation of high-stakes speaking tests. This knowledge can be further applied to develop scoring criteria for L2 oral assessment. Finally, future research in L2 pronunciation may provide new ways to operationalize stress or pitch as target variables to measure L2 speakers' oral production.

NOTE

The current study was presented at Pronunciation in Second Language Learning and Teaching 2012 under the title of *Pronunciation Features Distinguishing Examinees' Oral Performances at Different Proficiency Levels.* It examined salient pronunciation features that could distinguish Common European Framework of References (CEFR) speaking levels (B1-C2) in Cambridge ESOL General English Examinations. The full study appears in Cambridge ESOL *Research Notes 52* (2013). A comprehensive discussion of

the project funded by the Cambridge ESOL Research Program can be found in manuscripts currently under review for publication in other venues. Due to the constrained copyright of the articles, in the current proceedings, the author briefly reports one aspect of the findings, focusing on the relative impact of a wide array of pronunciation features on NNSs' oral proficiency.

ABOUT THE AUTHOR

Okim Kang is assistant professor of applied linguistics at Northern Arizona University. Her research concerns aspects of L2 pronunciation, speech perception and production, automated speech scoring, oral language proficiency assessment, language attitudes, and World Englishes. Email: okim.kang@nau.edu

Department of English
Liberal Arts Building 18
Room 140
Northern Arizona University
Flagstaff, Arizona, 86011-6032

Phone: 928.523.2059
Fax: 928.523.7074

REFERENCES

Brazil, D. (1997). The communicative value of intonation in English. Cambridge, UK: Cambridge University Press.

Fayer, J.M. & Krasinski, E.K. (1987) Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37*, 313–26.

Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing, 27,* 379-399.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct?. *Applied Linguistics, 29*(1), 24-49.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of l2 comprehensibility and accentedness. *System, 38,* 301-315.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal, 94*, 554-566.

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145-164.

Munro, M. J., & Derwing, T. M., (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech: The role of speaking rate. *Studies of Second Language Acquisition, 23*, 451–468.

Munro, M. J. & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System*, *34*, 520-531.

Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly, 35*(2), 233-255.

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language And Cognition, 15*(4), 905-916.

Wennerstrom, A. (1994). Intonational meaning in English discourse: A study of nonnative speakers. *Applied Linguistics* 15: 399-421.

Zechner, K., Higgins, D., Xi, X., Williamson, D. M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. *Speech Communication, 51*, 883-895.

VERSANT AND ADVANCED L2 SPEAKERS' RATINGS OF JAPANESE LEARNERS' ORAL ENGLISH

Akiko Okamura, Takasaki City University of Economics, Japan

It is important to evaluate the intelligibility of L2 speakers' English from a lingua-franca perspective. It is also now possible for computer programs to provide automatic evaluation of spoken English. It is of interest to know how such evaluations relate to those of L2 human listeners.

This study compared speaking test scores of intermediate proficiency level Japanese students generated by the machine-evaluated speaking test *Versant* with evaluations of the same informants made by advanced L2 speakers. Four Japanese informants' recordings were chosen such that overall *Versant* scores were similar but part-skill scores for 'fluency', 'pronunciation', 'grammar', and 'vocabulary' were different. Twenty-one Stockholm University English major students ranked the recordings twice, once in terms of intelligibility and once in terms of pronunciation quality. They also chose either segmentals or suprasegmentals as needing the most improvement.

The results show that what the program calls pronunciation is important for intelligibility. However, vocabulary seems also to affect intelligibility. Informants with lower *Versant* fluency scores were perceived as needing to learn suprasegmentals more than segmentals to improve intelligibility, while more fluent learners were perceived as needing to work on segmentals.

INTRODUCTION

Achieving intelligibility and comprehensibility is a priority for learners of English within the framework of World Englishes (Crystal, 2003; Jenkins, 2002, 2007; Kachru, 1992). According to Smith & Nelson (1985), intelligibility is the ability of the listener to recognize individual words or utterances while comprehensibility is the listener's ability to understand the meaning of the word or utterance in its given context.

Second language acquisition (SLA) research into intelligibility and comprehensibility has primarily looked at native listeners and non-native speakers of English (Derwing, Munro, & Thomson, 2008; Field, 2005; Hahn, 2004; Munro & Derwing, 1995). Both of these dimensions are usually connected to pronunciation and tend to be analyzed at the word or sentence level or are based on reading aloud (Derwing, 2003; Derwing & Munro, 1997, 2005; Derwing & Rossiter, 2003; Field, 2005).

The native listener/nonnative speaker model is not the only model that is relevant to speaking assessment, however. As the number of L2 speakers outnumbers that of L1 speakers, English has become a lingua franca for interaction between nonnative speakers around the globe, in what has come to be known as world Englishes (Crystal, 2003; Jenkins, 2002; Kachru, 1985, 1992). Kachru's (1985) three circles of world Englishes reflects the status and function of English in

different parts of the world. The inner circle refers to the countries where English is used as a mother tongue (i.e., the native speakers), while the outer circle can be used for countries where English has an official role and is used as a second language within the country. The expanding circle covers a wide range of countries where English is taught as a foreign language (i.e., the nonnative speakers). To show the interaction patterns of people in these three circles, Levis (2005) presented the speaker-listener matrix in Figure 1.

<div align="center">LISTENER</div>

		Inner Circle (IC)	Outer Circle (OC)	Expanding Circle (EC)
	Inner Circle	*IC–IC* *(NS-NS)*	**1. IC–OC**	*IC–EC* *(NS–NNS)*
SPEAKER	Outer Circle	**2. OC–IC**	**3. OC–OC**	**4. OC–EC**
	Expanding Circle	*EC–IC* *(NNS–NS)*	**5. EC–OC**	*EC–EC* *(NNS–NNS)*

Figure 1. World Englishes speaker-listener intelligibility matrix (from Levis, 2005, p. 373)

According to Levis (2005), research on intelligibility and comprehensibility is especially lacking in the bolded cells, where outer circle and expanding circle speakers/listeners interact (Jenkins, 2000, 2002, 2005; Kachru, 1992; Seidlhofer, 2003). There is great variation in terms of exposure to oral English in these areas of interaction. For example, speaking is the main focus in English classes in Sweden (Hincks, 2005), while in Japan grammar translation is still prevalent in high school English classes (Kikuchi, 2009).

The most noticeable difficulties connected to loss of intelligibility often come from L1 interference with L2 pronunciation (Flege & Freida, 1995; Flege, Munro & MacKay, 1995; Flege, Frieda & Nozawa, 1997a; Thompson, 2001). Intelligibility may also be an issue with stress and intonation (Celce-Murcia, Brinton & Goodwin, 1996; Morley, 1991). Prosodic errors may be more serious than segmental errors for non-English speakers (Munro & Derwing, 1999). Hahn's (2004) study showed that sentence stress errors have a negative impact on intelligibility. Field (2005) investigated recognition of misstressed English words and found that certain misstressed patterns are more serious than others. Analyzing suprasegmentals at a discourse level, Kang et al. (2010) showed that suprasegmentals accounted for nearly 50% of the variance in intelligibility scores.

However, the role of suprasegmentals for world Englishes may be more complex than previously thought. First, some research has found that native English speakers pay more attention to

individual sounds than suprasegmentals (Kondo, 2009; Riney, Takagi & Inutsuka, 2005). Second, most suprasegmentals may not be essential for communication among L2 speakers (Jenkins, 2006, 2007; Seidlhofer, 2003, 2004). Indeed, the importance of suprasegmentals may be different due to the type of L2 speakers involved in spoken interaction and their proficiency levels.

Although studies on intelligibility often focus on pronunciation of words or reading sentences, in naturally occurring English it is not easy to tease out pronunciation problems from other issues such as vocabulary and grammar (Morell, 2004; Pickering, 2004, 2006). Language testing often evaluates speaking skills using separate criteria such as grammar, vocabulary, fluency and pronunciation (Brown & Abeywichrama, 2010). However, the main concern of the field has been to maintain reliability and validity rather than to test intelligibility and comprehensibility (Davies, 2008; Meierkord, 2004).

Currently it has become possible for internet based computer programs to provide automatic evaluation of spoken English, and given the advantages in speed and expense of automated evaluation, it is of interest to know how such evaluations relate to those of L2 human listeners who are not native speakers of English. Thus, the use of an automatic evaluation can be better compared to the reality of English in today's world.

This study compared speaking test scores of intermediate proficiency level Japanese informants as generated by a language evaluation program (*Versant*) with evaluations of the same informants made by advanced L2 speakers. There are three research questions for the study.

Research questions

1. What is the relationship between L2 speakers' evaluation and the internet based computer program *Versant*?

2. What relationship exists between the perception of intelligibility and the linguistic criteria of proficiency test?

3. Is there any relationship between errors in segmentals and suprasegmentals, and the linguistic criteria of proficiency test?

Data collection

This study uses a computerized speaking test, *Versant,* to evaluate learners' speaking skills according to four criteria: *grammar* (i.e., sentence level syntax), *vocabulary*, *fluency* and *pronunciation* with overall scores in a range of 20 to 80. (Note: The average scores for Japanese college students tend to be between 30 to 40 according to the Versant Japan office.)
The students were given six spoken tasks of increasing difficulty over 17 minutes.

1. Read sentences,

2. Repeat sentences,

3. Answer questions,

4. Sentence building,

5. Story telling,

6. Open-ended questions

This study used the recordings of the 4 Japanese informants' responses to Question 6 (2

open-ended questions at 40 seconds per question (Appendix 1) for the evaluation because it was the only recording available on the internet. It was not evaluated automatically but was recorded for the identification of the examinee. Thus the part that the machine program evaluated was not the same as that for the Stockholm university students' evaluation. However, as the recording was part of the test, the human evaluation was conducted on the most difficult question.

The Japanese informants and the selection of the test recordings

There were 541 Japanese informants who took the test over the phone to measure their speaking skills in English at a regional university in Japan. They represented one expanding circle speakers of English. Four Japanese informants' recordings were chosen from the 541 informants so that overall scores were in the same range (41 to 46) but part-skill scores were different for *fluency* (42 to 50), *pronunciation* (39 to 53), *grammar* (31 to 49), and *vocabulary* (27 to 49) (see Table 1).

The L2 evaluators: 21 Stockholm university students

The L2 evaluators were 21 advanced learners of English who were taking a Second Language Acquisition course conducted in English at Stockholm University. They were chosen to represent one group of L2 speakers in the expanding circle, those who use English as a means of instruction at a university. The evaluators' nationalities were 12 Swedish, 3 Hong Kong Chinese, 2 German, 1 Austrian, 1 Finnish, 1 Korean, and 1 Vietnamese. They were asked to evaluate the Japanese informants' spoken English as part of their course work to understand the English learners' pronunciation problems.

The evaluation procedure

The recordings of the four Japanese informants' response to question 6 were downloaded from the internet and uploaded to a course website at Stockholm University. The evaluators were asked to listen to them as many times as they needed over two weeks to 1) rank the English in terms of intelligibility and pronunciation, 2) comment on English problems, and 3) evaluate how segments and/or suprasegments disrupted intelligibility (Appendix 2).

RESULTS AND DISCUSSION

The evaluators' assessment and the Versant test scores

The rankings of the four Japanese informants were converted into numbers on an interval scale from 1-4. Figure 2 shows that the evaluators' assessment on the bar graph, while the *Versant* scores were on the line graph with their overall scores at the bottom. The evaluators ranked the Japanese informants in the order of D, B, C, and A in both intelligibility and pronunciation, supporting a close relationship between intelligibility and pronunciation (Jenkins, 2000; Meierkord, 2004).

Notes: bar graph= Stockholm students' evaluation
line graph= the Versant test scores
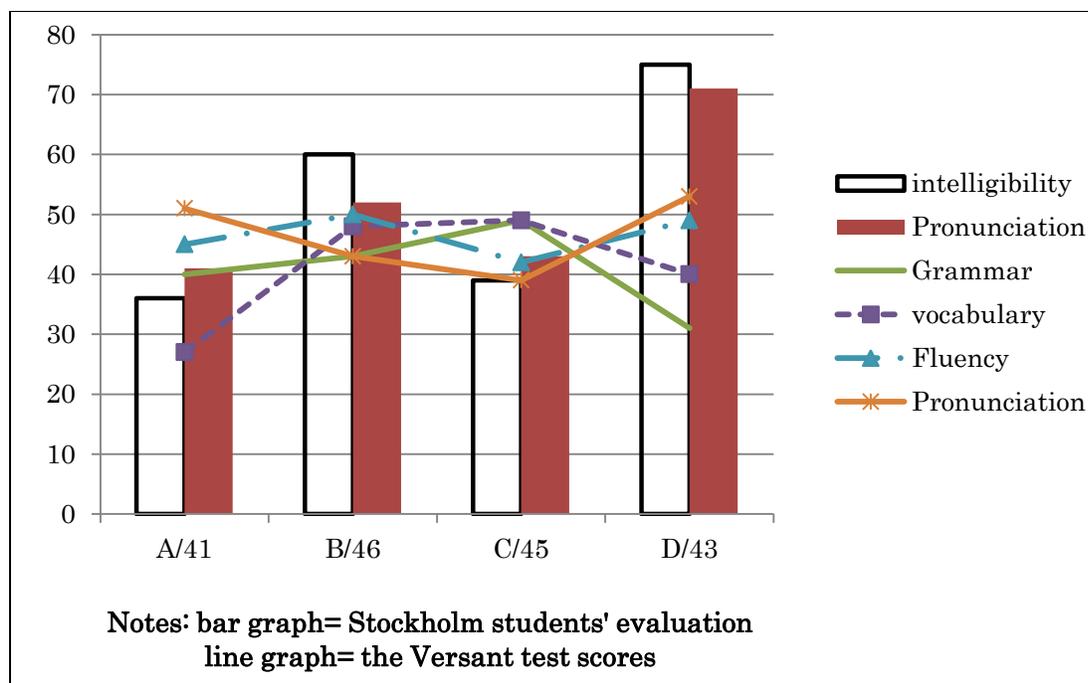
Figure 2. Versant scores of the Japanese informants ABCD and their ranking by Stockholm university students

The comments on A, B, C and D by the evaluators shown in Table 1 also strengthen the role of pronunciation in relation to other criteria. For example, in the *Versant* test the student C was the best in both *grammar* and *vocabulary*, but the worst in *pronunciation*.

Table 1

Versant Test Scores and Comments on the Intelligibility by Stockholm University Students

	Versant scores					Types of and number of comments from Stockholm students
	Overall	S G	Vocab	Flu	Pron	
A	41	40	**27**	45	51	Lack of vocabulary 3, lack of connected speech 1, Basic grammar problem 1, rhythm 1
B	46	43	**48**	50	43	Good vocabulary 3, better than A due to pronunciation and vocabulary, Good grammar but lack of accuracy 2, fluency OK 1
C	45	**49**	49	42	**39**	Strong Japanese accent 3, better grammar 2, bad grammar 2, vocabulary OK 2, limited vocabulary 1, fluency OK
D	43	**31**	40	49	**53**	Bad grammar 3, good intonation 2, good vocabulary 2, lack of vocabulary 1, fluency OK1, good accent 1

Notes: SG: Sentence grammar, Flu: Fluency, Pron: Pronunciation

All in all, as shown in Table 1, the evaluators' comments often reflected the *Versant* test scores. For example, in vocabulary in *Versant,* B gained higher scores than D and the evaluators commented *good vocabulary* (3 times) on B. Also on D, who was the best in pronunciation but worst in grammar in *Versant* had comments such as *good intonation* (twice) and *bad grammar* (3 times) by the evaluators. Overall, the L2 speakers' evaluations are not far from that of the

computer program based on the L1 norms.

Table 1 also shows that *vocabulary* is connected to judgments of intelligibility and suggests a further link to *fluency*. The evaluators stated that a large vocabulary size enables the learner to paraphrase the message, thus clarifying the meaning and avoiding pauses. The second best informant in pronunciation, A, was ranked third in intelligibility by the evaluators.

Poor *Versant grammar* scores, however, did not seem to reduce judgments of intelligibility as much as shown by Meierkord (2004). Informant D was ranked the highest in intelligibility by the evaluators (and had the highest *pronunciation* scores) but had lower *grammar* scores.

Comments on Japanese informants' pronunciation

Comments were given mainly on consonants (*see* Appendix 3). Those on the vowels were limited to /i/ and /o/ with both referring to the pronunciation of informants B and D, but only /i/ to that of A, and none for C. Only D, the strongest in pronunciation had comments on vowels from three evaluators. One stated that the student D had problems with vowels more than consonants. It seems that comments on vowels came only when the pronunciation was good enough to identify the problem as was also found among the comments on Japanese informants' English provided by English speaking teachers and non-teachers (Okamura, 2011).

A further analysis was conducted on the consonants, as consonants drew much more comments than vowels. Although the evaluators noticed problems with 14 types of consonants in total, the four Japanese informants A, B, C, D only received two types, /ɹ, l/ and /θ/ from a multiple number of the evaluators. As was shown by native speaking teachers and students on Japanese learners of English (Okamura, 2011), the most common problem seems to be /l/ and /ɹ/ identified by the evaluators (A: 9, B: 13, C: 6, D: 7). Because other sound problems were also related to the mother tongue such as /t/, /d/, /n/, /s/, and consonant clusters (see Thompson 2001), L1 transfer seems to be the major problem (Flege, 1995, 1997). /θ/ was also noticed by multiple evaluators (A: 1, B: 3, C: 5, D: 1). /θ/ was not included in the Lingua Franca core (Jenkins, 2000) because it was not considered to cause misunderstanding. However, unlike the voiced counterparts, voiceless sound of /θ/ may cause some misunderstanding as it can create minimal pairs such as *think/sink*, *worth/worse*, *thick/sick*. Moreover, it can be said that listeners are not always focusing on the intelligibility of the talk of L2 speakers as was found in social psychology and communication studies (Hosoda, Stone-Romero & Walter, 2007; Hosoda & Stone-Romero, 2009; Gluszek & Dovidio, 2010a, b; Gluszek, Newheiser & Dovidio, 2011).

Regarding the suprasegmentals, all the informants except D received "choppy" as the description of their English from the evaluators. The weakest informant in pronunciation, C was referred to as having a "strong Japanese accent" three times, suggesting a problem of prosody rather than just segmental sounds (since all the Japanese informants had similar segmental difficulties). Furthermore, as C received comments to slow down from eight evaluators, lack of English rhythm may have been interpreted as excessive speed. Choppy English can be intelligible but it would demand an extra effort on the listener, which can be part of the reason for some negative attitude to heavily accented English (Mugglestone, 1995). Thus learning the suprasegmentals may lead to a more positive evaluation. One evaluator wrote about D, who was the best in perceived pronunciation, "Thanks to his reasonably good accent and intonation, he had less serious problems than the rest."

The importance of improving segmental and suprasegmental sounds

Figure 2 shows the evaluators' choices about the importance of segmental and suprasegmental sounds for the four Japanese informants.
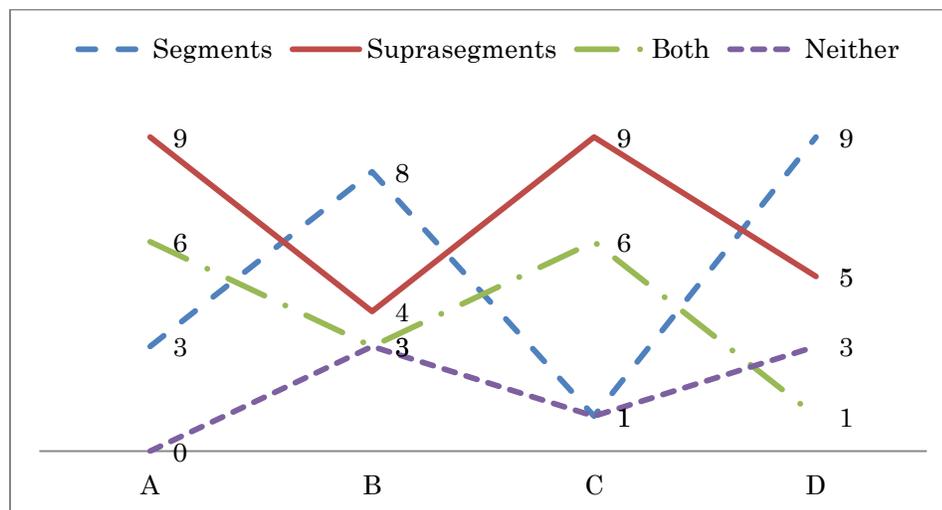
Figure 2. Importance of Segments and Suprasegmentals

Suprasegmentals were heard as being more important for A and C than B and D. It is interesting to note that A and C were given lower scores in fluency by *Versant* and intelligibility by the evaluators than B and D. Less fluent English appears to result in a perception of needing better suprasegmentals. As suggested by Kang et al. (2010), suprasegmentals seem to be highly related to intelligibility, which is also in agreement with the findings by Field (2005) and Hahn (2004). However, placing an emphasis on individual segmental sounds was suggested for those with higher rankings in intelligibility.

CONCLUSION

This study attempts to investigate what Japanese learners of English need to achieve intelligibility in the expanding circle interaction through the comparison of human and machine evaluation. This study showed that what the program called *pronunciation* was important for perceived intelligibility (Jenkins, 2002). The Japanese informant perceived as best in both intelligibility and pronunciation was also the one with the highest *Versant* scores in *pronunciation*.

However, this study also showed that pronunciation should be supported by vocabulary to increase intelligibility for L2 speakers as they can avoid the pronunciation problems through paraphrasing. By the same token, good pronunciation with lack of vocabulary seems to lead to unnecessary pauses, which can cause negative impressions of the speaker.

Informants with lower *Versant fluency* scores were perceived as needing to learn suprasegmentals more than segmentals to improve intelligibility, while more fluent learners were perceived as needing to work on segmentals. The *fluency* measure seemed to be associated with appropriate suprasegmentals, resulting in improved intelligibility. Machine evaluations can be usefully mapped on to listener perceptions, but interpretation is essential.

ABOUT THE AUTHOR

Akiko Okamura teaches English at Takasaki City University of Economics, Japan.   She obtained her PhD at Newcastle University, UK where she was a lecturer in Japanese. Her interests include cross-cultural communication, academic writing, and evaluation of L2 speakers' oral English.

REFERENCES

Abeywickrama, P., & Brown, H.D. (2010). *Language assessment: Principles and classroom practices*, 2nd ed. Pearson Longman.

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages.* New York: Cambridge University Press.

Crystal, D. (2003). *English as a global language.* Cambridge: Cambridge University Press.

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25,* 327-347.

Derwing, T.M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review, 59,* 545-564.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. T*ESOL Quarterly, 39*, 379-397.

Derwing, T. M., & Munro M. J. (1997). Accent, comprehensibility and intelligibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19, 1-16.*

Derwing, T., & Rossiter, M. (2003). The effect of pronunciation instruction on the accuracy, fluency and complexity of L2 accented speech. *Applied Language Learning, 13,* 1–17.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics, 29*, 359–380.

Deterding, D., & Kirkpatrick, A. (2006). Emerging southeast Asian Englishes and intelligibility. *World Englishes 25*, 391−409.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39,* 399–423.

Flege, J., & Freida, E. (1995). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25,* 169–186.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America,* 97, 3125–3134.

Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics, 25*, 169–186.

Gluszek, A., & Dovidio, J. F. (2010a). Speaking with a nonnative accent: Perceptions of bias, communication difficulties, and belonging in the United States. *Journal of Language and Social Psychology*, 29, 224-234.

Gluszek, A., & Dovidio, J. F. (2010b). The way *they* speak: A social psychological perspective on the stigma of non-native accents in communication. *Personality and Social Psychology Review*, *14*, 214-237.

Gluszek, A. Newheiser, A.-K., & Dovidio, J. F. (2011). Social psychological orientations and accent strength. *Journal of Language and Social Psychology*, *30*, 28–45.

Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38,* 201–223.

Hincks, R. (2005). Measures and perceptions of liveliness in student oral presentation speech: A proposal for an automatic feedback mechanism. *System*, *33*, 575-591.

Hosoda, M., Stone-Romero, E. F., & Walter, J. N. (2007). Listeners' cognitive and affective reactions to English speakers with standard American English and Asian accents. *Perceptual and Motor Skills*, *104*, 307-326.

Hosoda, M., & Stone-Romero, E. (2009). The effects of foreign accents on employment related decisions. *Journal of Managerial Psychology*, *25*(2), 113-132.

Jenkins, J. (2000). *The phonology of English as an international language*: New Models, New Norms, New Goals. Oxford: Oxford University Press.

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

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

Jenkins, J. (2006). Current perspectives on teaching world Englishes and English as a lingua franca. *TESOL Quarterly, 40*, 157-81.

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity*. Oxford: Oxford University Press.

Kikuchi, K. (2009). Listening to our learners' voices: What demotivates Japanese high school students? *Language Teaching Research, 13*, 453-471.

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson *(Eds.), English in the World: teaching and learning the language and literatures* (pp. 241-255). Cambridge: Cambridge University Press for The British Council.

Kachru, B. B. (1992). Models for non-native Englishes. In B. B. Kachru (Ed.), *The Other Tongue: English Across Cultures* (pp. 48-74). Urbana: University of Illinois Press.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*, 554-566.

Kondo, M. (2009). Is acquisition of L2 phonemes difficult? Production of English stress by Japanese speakers. In *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA 2009)*, (pp. 105-112). Somerville, MA: Cascadilla Proceedings Project.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39,* 369–377.

Meierkord, C. (2004). Syntactic variation in interactions across international Englishes. *English World-Wide*, 25, 109–132.

Morell, T. (2004). Interactive lecture discourse for university EFL students. *English for Specific Purposes,* 23, 325–338.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other

languages. *TESOL Quarterly, 25,* 481–520.

Mugglestone, L. (1995). *'Talking proper': The rise of accent as social symbol.* Oxford, England: Clarendon.

Munro, M., & Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.

Munro, M., & Derwing, T. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 49 (Supp. 1),* 285–310.

Okamura, A. (2011). What are the specific barriers to intelligibility for native English speakers when listening to oral presentations delivered by Japanese speakers of English? Paper presented at 5th International Conference on Native and Non-native Accents of English (*Accents 2011,* Łódź, Poland). December 2011.

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes,* 23, 19–43.

Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics, 26*, 219–33.

Riney, T.J., Takagi, N., & Inutsuka, K. (2005). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly, 39*, 441-466.

Seidlhofer, B. (Ed.) (2003). *Controversies in applied linguistics*. Oxford: Oxford University Press.

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics*, *24*, 209 – 242.

Smith, L., & Nelson, C. (1985). International intelligibility of English: Directions and resources. *World Englishes, 4,* 333–342.

Thompson, I. (2001). Japanese speakers. In M. Swan & B. Smith (Eds.), *Learning English*, 2nd ed. (pp. 296-309). Cambridge: Cambridge University Press.

Appendix 1

1. Versant Questions 6 as answered by the four Japanese informants

Japanese student A

1 Is it best for family members to help children with school work, or to let children complete school work alone without help? Please explain.

2 After the school day is over, some families allow children to play as long as they wish, while other families believe children should continue studying throughout most of the evening. What do you think is a good method for determining how much leisure time a child should have?

Japanese student B

1 In your opinion, where is the best location to raise a family? Why do you feel this way?

2 From your point of view, is it better to work in a large organization or in a small organization? Please explain your opinion.

Japanese student C

1 Today, many people are living longer, and many old people are being cared for by their children or grandchildren. How does this situation affect the family?

2 Should parents choose the course of study for their children, or should children choose their own course of study? Please explain your thinking.

Japanese student D

1 Some people think that week-ends should be spent resting and relaxing, while other people prefer to use the extra time to complete work. How do you like to spend week-end time? Please explain.

2 Do you think that every person in the family should help with household tasks, even very young children? Please explain your thinking.

Appendix 2

2. Swedish students' evaluation of Japanese learners' English

2.1. The first round: ranking in intelligibility and pronunciation

You are going to listen to four Japanese informants responding to questions in English. Please rank these four in terms of intelligibility and pronunciation.

Intelligibility and pronunciation ranking

	Speaker 1	Speaker 2	Speaker 3	Speaker 4
Intelligibility ranking				
Pronunciation ranking				

2.2. The second round: segmentals and suprasegmentals

You are going to listen to the same speech again. You may find that intelligibility difficulties arise from both individual sound problems and sentence level problems.

2.2.1. Which interfered more with your understanding?

(   ) A. Pronunciation of certain sounds such as /r/ or /s/

(   ) B. Pronunciation at sentence level such as English rhythm or English intonation.

(   ) Both A and B

(   ) Neither A nor B

2.2.2. What do you think are the most serious problems for them in terms of making their speech intelligible?

Pronunciation: segmental and suprasegmentals

Name	A, B, Both AB, Neither A nor B	Their difficulties
Speaker 1		
Speaker 2		
Speaker 3		
Speaker 4		

Appendix 3

Number of comments about segmental and suprasegmentals pronunciation problems

Category	A	B	C	D
Consonants	6 types/ 14 tokens	4 types/ 19 tokens	5 types/ 15 tokens	7 types/ 13 tokens
Vowels	1 type/1 /i/	2 types/1 /i/,/o/	0 type/	2 types/3 /i /, /o/
Supra-segmentals	3 types/8 Lack of Stressed syllables (2), Choppy rhythm (3) Intonation and fluency (3)	1type/2 Choppy rhythm (2)	1 type/ 3 Strong Japanese accent/ Rhythm (3)	1 type/1 Choppy rhythm (1)
Perceived speed	too slow: 1 too much hesitation: 7	fast : 1 long pause: 2	too fast: 8	Short pause : 3

TEACHERS' VIEWS ON THEIR PROFESSIONAL TRAINING AND ASSESSMENT PRACTICES: SELECTED RESULTS FROM THE ENGLISH PRONUNCIATION TEACHING IN EUROPE SURVEY

[1]Anastazija Kirkova-Naskova, University of Skopje

Elina Tergujeff, University of Jyväskylä

Dan Frost, Université de Savoie

Alice Henderson, Université de Savoie

Alexander Kautzsch, University of Regensburg

David Levey, University of Cádiz

Deirdre Murphy, Trinity College Dublin

Ewa Waniek-Klimczak, University of Łódź

The English Pronunciation Teaching in Europe Survey (EPTiES) is a collaborative effort by a group of European researchers interested in the state of English pronunciation teaching in Europe. Given the lack of research-based information on the topic (cf. e.g. Foote et al. 2011, Macdonald 2002), ten researchers designed an extensive online survey, which attracted participants from all over Europe. The participants are EFL/ESL teachers from various teaching contexts. This paper concentrates on two parts of the survey that deal with teacher training and pronunciation assessment. Responses from seven European countries are analysed (n=630). We present findings concerning the contents of teacher training received by the respondents and their overall evaluation of it. In addition, we look into the respondents' pronunciation assessment methods.

INTRODUCTION

In the past two decades, a number of studies have focused on English pronunciation teaching and it is evident that there is a growing interest in this field of research. Various attempts have been made to deal with relevant issues related to teaching practices, materials, training and attitudes to native speaker models both from the teachers' and the learners' perspective. Most of these studies have been conducted in ESL settings and in English-speaking countries such as Canada (Breitkreutz et al., 2001, Foote et al., 2011), the USA (Murphy, 1997), Australia (Macdonald, 2002), and Great Britain (Bradford & Kenworthy, 1991, Burgess & Spencer, 2000). Pronunciation teaching in contexts where English is taught as a foreign language (EFL) has been studied for example in Spain (Walker, 1999), Finland (Tergujeff, 2012), and in EFL environments of Ireland (Murphy, 2011). In addition, attitudes towards native speaker models and the degree of success in reaching the models in an EFL context have been investigated in Poland (Nowacka, 2010; Waniek-Klimczak, 2002; Waniek-Klimczak & Klimczak, 2005), Serbia (Paunović, 2009), and Bulgaria (Dimitrova & Chernogorova, 2012). It appears that recent

---

[1] Anastazija Kirkova-Naskova and Elina Tergujeff share the responsibility over the present paper. The rest of the authors - listed alphabetically - have contributed in the following stages of the research: survey design, data collection, analysis, and commenting on the manuscript. The EPTiES research group is led by Dr Alice Henderson.

research has given the appropriate attention to important aspects of pronunciation teaching; they are, however, looked into separately and tend to be rather country-specific with comparative studies, particularly between European countries, few and far between. Given this lack of research, the English Pronunciation Teaching in Europe Survey (EPTiES) seeks to provide a detailed insight into the current state of affairs. It addresses the issue of pronunciation teaching within an EFL setting in various European countries from a teachers' perspective.

Initial results rendered from a fraction of the EPTiES data came from the analysis carried out by Henderson and her team of researchers (Henderson et al., 2012). Findings revealed that the respondents feel that English is a very important language in relation to other languages, and consider pronunciation, in particular, to be one of the most important language skills. They further rate their own pronunciation skills as very good and claim to have a relatively high awareness of their learners' goals and skills in English pronunciation. According to the teachers, their learners strive for a native-like pronunciation to some extent, and generally prefer General American (GA) as a pronunciation model. On the other hand, the teachers themselves demonstrate a tendency to use British Received Pronunciation (RP) in their teaching.

In this paper, we further analyse two subsets of data from EPTiES addressing areas of pronunciation pedagogy related to professional teacher training and practical approaches to pronunciation assessment. The aim is primarily to tackle issues relatively underrepresented in the literature and to explore recurring trends in the European context.

The question of the lack of teacher training deserves attention and has been raised in several studies. Surveys conducted in Canada (Breitkreutz et al., 2001, Foote et al., 2011), Australia (Macdonald, 2002) and the USA (Murphy, 1997) suggest that many teachers teach pronunciation without substantial pedagogical training in this area, and that they often wish that they had received a more extensive training. Our study presents a closer look at what EFL teachers in Europe think about their teacher training relevant to pronunciation teaching with emphasis on the amount and contents of the training.

Pronunciation assessment is yet another challenge that teachers are faced with in their professional lives. Despite the fact that pronunciation is a vital component of proficiency in spoken English, little published work seems to exist which addresses the issue of pronunciation testing and evaluation. This absence can be accounted for partly by the fact that "precise identification of pronunciation problems can be difficult even for experienced listeners" (Yates, Zielinski, & Pryor, 2011, p.4) and also by the fact that "the large body of literature on language assessment applies to pronunciation just as it does to any skill - reading, listening, speaking" (Celce-Murcia et al., 1996, p.341). Studies that have investigated issues related to pronunciation assessment mainly focus on the reliability of the descriptors for the speaking part in standardised tests such as Cambridge ESOL exams and IELTS (Brown & Taylor, 2006; DeVelle, 2008; Hubbard, Gilbert & Pidcock, 2006; Szpyra-Kozłowska et al., 2005; Yates, Zielinski & Pryor, 2011). Researchers have also been interested in what aspects of pronunciation assessment examiners should refer to while assessing individual speakers' speech and/or spoken interaction. With this in mind, research has addressed issues such as the relevance of diagnostic assessment versus holistic/impressionistic/global assessment (Levis, 2006; Szpyra-Kozłowska et al., 2005), the importance of intelligibility and comprehensibility (Gass & Varonis, 1984; Jenkins, 2000; Munro & Derwing, 1999), as well as the intricacies involved in assessing accuracy and fluency (Kormos & Dénes, 2004; Levis, 2006). This study attempts to broaden the research scope by providing insights into the use of established reference scales such as the Common European

Framework of Reference (CEFR, Council of Europe, 2001), which has been promoted in the evaluation of language skills throughout Europe. The present article also explores what types of popular classroom activities are used as pronunciation assessment instruments.

THE SURVEY: METHOD AND PARTICIPANTS

EPTiES is an ongoing collaborative research project with partners from various European universities, the aim of which is to compare and contrast the English pronunciation teaching practices and attitudes of English language teachers working in EFL contexts across Europe. The data come from an online survey of 843 teachers from 31 European countries carried out between February 2010 and September 2011. Most of the respondents were female (76%) with an average age of 43 years. The majority were non-native speakers (89%), predominantly working in the public sector (85%) with 15 years teaching experience on average.

The survey consisted of 57 questions grouped into nine sections[2]: participant information, outside classroom, pronunciation teaching methods, teaching materials, evaluation of pronunciation, teacher training, views/attitudes, and model/norm. Some of the questions were formulated to reflect specific national contexts. Both closed-ended and open-ended questions were incorporated in the survey. The following types of closed-ended questions were used: Likert-scales, yes-no items, and multiple-choice items. The Likert-scales were of the five-point format with descriptive terms as response options adapted to the relevance of the question, for instance, in teachers' evaluations of their own pronunciation skills (with 1 as "extremely poor" and 5 as "excellent"). Where more straightforward answers were required, yes-no items were used with an additional option for a clarification comment. In all sections of the questionnaire, multiple-choice questions with the option of choosing more than one answer were used (for example in the section on model/norm for mapping the responses on productive and receptive work). In the analysis, frequencies were calculated for the quantitative data, whereas the open-ended questions were analysed qualitatively by coding the answers for recurrent themes (Dörnyei & Taguchi, 2009).

In the present paper, a more limited data sample (n=630) is used consisting of the respondents from seven countries where a minimum of ten people responded. In alphabetical order the countries are: Finland, France, Germany, Ireland, Macedonia, Poland and Spain, and the number of respondents per country is presented in Table 1. Respondents from these countries form the majority of the responses in total. Not all of the respondents completed the survey, but all responses are considered for those parts of the questionnaire that were filled in.

---

[2] Results from teacher training, views/attitudes and model/norm sections were presented in Henderson et al. (2011) and reported in Henderson et al. (2012). Results related to pronunciation teaching methods, teaching materials and outside classroom were presented in Henderson et al. (2013a) and reported in Henderson et al. (2013b).

Table 1
Number of respondents per country

Country	Number of respondents
Finland	103
France	65
Germany	363
Ireland	12
Macedonia	36
Poland	20
Spain	31
Total	630

The demographic data reveals that the present sample is well representative of the whole survey. The respondents were predominantly female (77%) with an average age of 43 years. The age range was from 21 to 69 years. Their teaching experience varied from 0 to 44 years with an average of 11 years. A vast majority reported working in the public sector (88%), and 92% of the respondents of the present sample were non-native speakers of English. This background information about the respondents is illustrated in Table 2 below.

Table 2
Background information about the respondents

gender (n=629)		age (n=630)		native speaker status (n=629)		teaching experience (n=628)		teaching context (n=629)	
female	**male**	**mean**	**range**	**NS**	**NNS**	**mean**	**range**	**public**	**private**
483 (77%)	146 (23%)	43.27 years	21–69 years	53 (8%)	576 (92%)	11 years	0–44 years	552 (88%)	77 (12%)

FINDINGS

Teacher training

The survey included a number of questions related to the teacher training the respondents had received. The respondents were initially asked to give an overall rating of their teacher training in relation to pronunciation on a scale from 1 to 5, with 1 as "extremely poor" and 5 as "excellent". The results show that the whole scale was used, with the average rating of 2.91 (n=478).

The respondents were then required to explain how much training they had received *specific to teaching English pronunciation* (see Table 3). This question was open-ended, and the qualitative content analysis of the responses revealed the following as recurrent themes in all seven countries: (1) phonetics/pronunciation courses/modules; (2) pronunciation component as part of a more general TEFL course or MA programme (except Spain); and (3) no or hardly any training. Some respondents from Finland, France and Germany regarded practical sessions with

native speakers to be part of their training, in addition to a stay in an English-speaking country for the purpose of taking a language course. A number of items arose sporadically in the qualitative data, for example additional training via conferences, seminars, summer schools and workshops or different time periods devoted to training with no specific description about the type of training received. These items varied so much that no generalisations could be made.

Table 3
Summary of the qualitative content analysis regarding amount of training
("Please tell us how much training you received specific to teaching pronunciation.")

Items mentioned	Total number of mentions	Countries where mentioned
Phonetics course as part of their undergraduate studies	135	FIN, FRA, GER, IRE, MAC, POL, SPA
Several pronunciation or phonetics courses (not specific)	57	FIN, FRA, GER, POL, SPA
Intertwined with other topics e.g. As part of a general TEFL course	31	FIN, FRA, GER, IRE, MAC, POL
Native speaker contact or a stay in an English-speaking country	37	FIN, FRA, GER
Hardly any/very little	45	FIN, FRA, GER, IRE, MAC, POL, SPA
None	51	FIN, FRA, GER, SPA
Don't remember	35	FIN, FRA, GER

Teachers' comments reveal that the phonetics/pronunciation courses were in most cases undergraduate courses or modules which did not aim at preparing the future teachers to teach pronunciation, but to improve their own pronunciation, bearing in mind that the majority of the respondents were non-native speakers of English. The number of courses varied from one (frequently mentioned) to four (rarely mentioned). The following representative responses summarize the situation: "Three or four courses of pronunciation, but its intention was to improve our pronunciation not to teach us to teach it. I've had to figure out myself how to do it" (#846), and "We had an exam in Phonetics at the University, which was great; but very little training in teaching pronunciation" (#657). As illustrated, many of the respondents clearly regarded their undergraduate courses in pronunciation and phonetics as part of their training, and described how they themselves had practiced their own pronunciation. Although these general degree courses undoubtedly provide an important foundation, it was striking how few respondents mentioned having received subsequent courses dealing specifically with ways to *teach* pronunciation.

It was evident from the comments that training in pronunciation pedagogy does not generally include a separate pronunciation-oriented course, but rather is made up of smaller intervals of theoretical lectures and/or pronunciation activities as part of more general TEFL courses or as part of MA studies. Respondents frequently indicated that their "… teacher training was an all-

round course with different aspects of teaching combined into a programme which included teaching pronunciation" (#857) or even too general as in "I have not received any training specific to teaching pronunciation, my training was rather global and pronunciation incorporated into it" (#726).

Unfortunately, respondents frequently mentioned that they had received no specific training in this area, and many reported small amounts of training in phonetics. Where there was no or little training provided, the teachers compensated with "…self-teaching and a good deal of experience" (#741).

With a survey of this scale and having in mind that teacher training in general entails training in assessment, one would expect respondents to provide details of their experience of how well they were prepared to assess pronunciation. To our disappointment, only two respondents referred to pronunciation assessment separately. The first respondent mentioned: "personal training and assessment of pronunciation at university (gave an idea of what it should be like) (#60). The second respondent indicated lack of training in this area: "Usually it comes down to teaching pronunciation exercises and ideas. Never assessment of pronunciation or systematic work with it" (#678).

In another question in the survey, respondents were asked to describe the contents and/or style of the training they received (see Table 4). This question was also open-ended thus a qualitative thematic analysis was carried out to group the responses in most recurring themes. Participants from Finland, France, Germany and Poland gave descriptions of what appeared to be training to improve their own pronunciation through practical classes in language labs. As reported in the responses, these mostly covered listening, reading aloud words/sentences/texts and phonetic/phonemic transcripts, as well as "… work on minimal pairs/repetitions…" (#826), "stress and intonation exercises" (#485), and "…exposure to a variety of accents…" (#708). In all seven countries the respondents were taught how to implement IPA symbols and had practical sessions in phonetic/phonemic transcription of sounds, words, utterances and texts. In addition, data from all seven countries showed that the respondents' training had a theoretical orientation with lectures in phonetics and phonology, revealing a frequent overlapping use of the terms pronunciation, phonetics and phonology. For example, in response to this question about their training to teach pronunciation, one respondent wrote that the university module "…was theoretically based without any practical classroom application. This gave me a good overview of the IPA and the different terms related to the biology of the mouth along with the restrictions some speakers may have" (#470).

Table 4
Summary of the qualitative content analysis regarding content/style of training
("Please explain the content and/or style of the training you received.")

Items mentioned	Total number of mentions	Countries where mentioned
Language lab	81	FIN, FRA, GER, POL
Transcription, Phonetic training, Implementing IPA symbols	82	FIN, FRA, GER, IRE, MAC, POL, SPA
Lectures/theory	44	FIN, FRA, GER, IRE, MAC, POL, SPA
Weekend seminars, In-service training, Seminars, Classes abroad	24	GER
Specific activities:		
- **Repetition/drills**	21	FIN, FRA, GER, POL
- **Conversation**	17	FIN, FRA, GER, POL
- **Listening tasks**	14	FIN, FRA, GER, MAC, POL
- **Reading aloud**	13	FIN, FRA, GER, POL
University classes to improve one's own pronunciation	105	GER
Learning by doing, Individual self-improvement	15	GER
None	33	GER, SPA, POL
Don't remember	29	FIN, FRA, GER, MAC, SPA

When it comes to the practical aspects of their training, the following specific activities were noted (see Table 4 for specific countries where mentioned): repetition/drills, conversation, listening tasks and reading aloud. One positive experience comes from several German respondents who seem to have improved their practical skills through in-service training and weekend seminars. For instance, one of them gave a detailed description of the type of training received: "An experienced teacher showed us her methods. We examined materials, tried it out as if we were the children and learned about studies in connection with the way children learn how to speak and understand English" (#378). Another respondent mentioned training practice with a native speaker: "A training day with a native speaker; the content was to train pronunciation and classroom methods" (#450). In this section of the survey, many German respondents quite specifically mentioned university classes as a means to improve their own pronunciation. Also, they referred to "learning by doing". This was not mentioned in any of the data from other countries.

Finally, 21 activities in total were mentioned between 1 and 5 times either in one country only or in several countries separately. They can be grouped as follows: (1) *teacher training specific activities* (e.g. watching other teachers, training in evaluation of oral skills, training in different

accents, stress on the importance of pronunciation teaching); (2) *teacher-student activities* (e.g. listening and correction, teacher correcting/commenting on students' pronunciation, identifying students' problems, L1 influence on L2 sounds, evaluating oral exams, oral translations, diagnostic tests of students' recordings at the beginning and the end of the year); (3) *student-student interactive classroom activities* (rhymes, games, stories, singing, film analysis, role-plays, pair work and group work, reading plays); and (4) *outside classroom activities* (recording one's own speech, exposure to English TV programmes).

Pronunciation assessment practices

Pronunciation assessment was approached in several questions in the survey. First, we inquired whether the teachers base their assessment on an established scale, either national or international. In our sample, to this question 497 respondents provided an answer, the vast majority of which (84.71%) replied that they did not, while the remainder (15.29%) answered affirmatively. Of those who indicated using an established scale (n=76), 85.53% referred to the Common European Framework of Reference (CEFR; Council of Europe, 2001).

We were also interested to find out *when* teachers assess their students' pronunciation. Results show that initial diagnostic assessment at the beginning of the course is practised by 31.06% (n=498), while the other 68.94% do not make use of this assessment technique. In contrast, teachers seem to prefer during-the-course assessment (48.80%), they rarely opt for end-of-course assessment (5.62%), or sometimes combine these two types (32.93%).

The practical side of pronunciation assessment was explored by including a set of questions about assessment approaches and tasks. The respondents were asked to indicate which of the listed tasks they use in assessing their learners' pronunciation separately for diagnostic, formative and evaluative assessment. The results are presented in Table 5.

Table 5
Tasks used in the assessment of pronunciation skills. (n=504)

Type of task	Diagnostic	Formative	Evaluative
Written work, e.g. transcription	8.53%	19.64%	19.64%
Oral performances, e.g. presentations	27.58%	80.36%	72.42%
Individual oral exams	14.68%	39.48%	44.64%
Oral exams in pairs	15.48%	47.42%	55.16%
Listening and questions	22.62%	67.86%	60.71%
Reading aloud	27.38%	75.60%	58.93%
Other	4.76%	12.10%	8.13%
I don't know	0.60%	1.39%	1.59%
None of the above	9.92%	5.36%	6.35%

In relation to types of assessment used, it is apparent that formative assessment is the preferred approach; all of the tasks are predominantly used for this type of assessment except for

individual exams and oral exams in pairs, which more teachers say they use for evaluative assessment. This set of data also points out the lesser-used method of diagnostic assessment.

When looking at the types of tasks used to assess pronunciation, respondents chose written tasks (e.g. transcription) the least frequently for all types of assessment (in all three cases below 20%). In contrast, oral performances are most frequently chosen for all types of assessment with the highest value for formative assessment (80.36%), followed by reading aloud (75.60%), listening and questions (67.86%), and oral exams in pairs (47.42%). Although respondents indicated that diagnostic assessment is not a common practice, it is no surprise to find that reading aloud is most frequently chosen to diagnose learners' pronunciation with almost one-third of the respondents choosing this option (27.38%).

DISCUSSION AND CONCLUSION

Based on the findings from the *teacher training* subset of data presented in this paper, we can conclude that the teachers in our sample are in general critical of the training they received. In addition, some of them make a distinction between the training they received to improve their own pronunciation and the training they received for teaching pronunciation as professionals. While they rate their own pronunciation quite high (Henderson et al., 2012), they mention a lack of training in teaching pronunciation. Not receiving the professional development they need, teachers may feel reluctant to teach pronunciation, as pointed out by Foote et al. (2011, p. 16). This may in some cases have an adverse effect on the quality of pronunciation teaching; in the worst-case scenario teachers may neglect pronunciation teaching completely. Given that EFL teachers in various European countries and ESL teachers in the USA, Canada and Australia feel they lack training in how to teach pronunciation, this would seem to be a global problem, leading us to reflect on how to improve teacher training programmes.

The analysis of the *evaluation of pronunciation* subset of data reveals that the respondents' classroom practices may not reflect the sorts of pronunciation assessment proposed in the pronunciation teaching texts and professional literature, at least not in all aspects. Diagnostic assessment, for instance, is a common type of testing favoured by many pronunciation experts. Celce-Murcia et al. (1996, p. 341) recommend it as "the teacher's initial method of setting or adjusting curricular objectives". In his recent publication on teaching English as a lingua franca, Walker (2010, p. 148) advocates the use of diagnostic tests to help teachers "find out what problems the learner has with the language". Moreover, Levis (2006, p. 247) discusses in favour of diagnostic assessment, arguing that teachers "need to become aware of the relevant phonological categories and be able to name important errors … [because] being able to diagnose pronunciation in detail makes the teachers more fit to assess standardized tests, which is important since teachers are the primary source of raters". Our results, however, show that teachers do not seem to be using diagnostic assessment frequently, which may imply that: (1) teachers are not prepared to perform such detailed feedback analysis of learners' pronunciation; (2) it is not required in the course curriculum as a testing method; or (3) they simply lack the time and/or technical resources to do so.

On the other hand, the other two types of assessment, formative and evaluative, are almost equally represented in our sample. It's not surprising that formative assessment is used as it typically takes place during learning, focuses on helping students with the learning process, and aims at improving learning (Huhta, 2010). Brown (2003) argues that formative assessment

provides washback in the form of information to the learners on their progress, thus adding an important dimension to the effect of testing on learning in this case. In light of the survey results, we consider the practice of formative assessment by our respondents to reflect a positive trend in the European context. It also supports the tendency demonstrated by our respondents to assess their students during the course. As for their use of evaluative assessment, based on the types of tasks that are used i.e. individual exams and oral exams in pairs, we may infer that this type of assessment is practiced when a holistic judgement of learner's pronunciation performance is required.

It is generally accepted that pronunciation poses the greatest difficulty for assessment because it incorporates both knowledge and skills components, and, at the same time, it is viewed as both receptive and productive skill. It is, therefore, expected that a range of tasks be employed as assessment instruments. While oral performances and reading aloud evaluate the productive skills in learners, listening and transcription evaluate their receptive skills. Our results show that participants claim to use the first two most frequently, followed by listening and questions but not transcription. The participants seem to be covering both the receptive and productive aspects of pronunciation assessment, in one form or another and at one point or another in a course.

It was surprising to find out that few teachers base their assessment on an established scale. This result could be viewed in the following context: the respondents are experienced teachers from various countries mainly working with monolingual classes and probably experiencing frequent pronunciation errors typical of speakers of certain L1s. Perhaps, this language-specific aspect should be incorporated in a more context-based pronunciation assessment scale if we consider how criticized CEFR is for being too general and too broad for application (Fulcher, 2004; Milanović, 2002; Weir, 2005). This would not be the first attempt to improve a scale, as some have already been made for the English language within the English Profile Programme (www.englishprofile.org).

In terms of teachers' preparation to assess pronunciation, our data certainly did not yield any relevant comments to draw conclusions from. However, the following noticeable mismatch should be pointed out: the majority of the teachers were trained in phonemic/phonetic transcription but do not make use of it in evaluating learners' performance. Based on data from another section in the EPTiES survey (Henderson et al., 2012) this may have to do with the influence of the Communicative Approach as promoted via the CEFR and other European policies: teachers may prefer to assess their students' pronunciation in situations closer to real-life communication than by phonemic/phonetic transcriptions.

As a follow-up to the insights about teachers' pronunciation assessment practices presented in this paper, further research is needed to find out why particular approaches are preferred and whether they are based on overall evaluation of pronunciation or on discrete items of pronunciation such as individual sounds, stress patterns, rhythm and intonation. Further inquiry might provide useful information about specific assessment tasks, such as which aspects of oral presentations are assessed and what features are rewarded and/or penalized.

To summarize, the aim of the present paper was to explore two aspects of English pronunciation teaching, in particular teacher training in teaching pronunciation and pronunciation assessment practices, viewed from the perspective of teachers who work in an EFL setting in various countries in Europe. With respect to teacher training, the results reveal that on average the teachers are not satisfied with their training in this area, and many are totally lacking in training.

With respect to pronunciation evaluation, it is evident that they assess their students during the course adopting a formative approach, and, surprisingly, few base the assessment on an established scale, such as the CEFR. Pronunciation assessment tasks seem to reflect the types of assessment they may have received in their schooling and written work is seldom used. In line with previous research, these findings confirm the apparent need for more structured teacher training programmes that assign a fair share of time to pronunciation instruction and assessment.

ACKNOWLEDGEMENTS

CORRESPONDING AUTHORS

Anastazija Kirkova-Naskova Email: akirkova@t-home.mk

Elina Tergujeff Email: elina.tergujeff@jyu.fi

ABOUT THE AUTHORS

Anastazija Kirkova-Naskova is a teaching assistant and researcher in English Phonetics and Morphology at the Department of English Language and Literature, University of Skopje, Republic of Macedonia. She is currently working on her PhD thesis in the area of speech perception and the impact of perceptual training on pronunciation proficiency of EFL learners. She is actively involved in two on-going research projects: *English Pronunciation Teaching in Europe Survey – EPTiES*, and *Corpus of Macedonian English Interlanguage* (national project, Ministry of Education and Science, Republic of Macedonia). Her research interests focus on pronunciation instruction, L2 phonological and prosodic acquisition, L1 interference and accent. She is also a trained UCLES oral examiner. Email: akirkova@t-home.mk

Elina Tergujeff is a professional EFL teacher and a teacher of Swedish as a second language. She has worked as junior researcher in the Department of Languages, University of Jyväskylä, Finland, 2010–2013. Tergujeff completed her PhD in 2013 on English pronunciation teaching in Finland. Her mixed methods approach resulted in a multi-perspective mapping of English pronunciation teaching practices in Finnish schools, ranging from primary to upper secondary level. She currently works as research coordinator of the University of Jyväskylä Language Campus. Email: elina.tergujeff@jyu.fi

Dan Frost studied languages and linguistics at the University of York in the UK and at Strasbourg, Aix-en-Provence in France. His doctorate is in English for Specific Purposes and Teaching Theory. He taught English in the UK, Thailand and Sweden before settling in France, where after teaching in secondary schools and then in ESP & EAP for 10 years in an IT department, he now lectures in the Applied Foreign Languages department of the University of Savoie, in Chambéry, France. He also works with trainee teachers, especially in the area of teaching pronunciation. His main research interests are teaching pronunciation, oral English and computer-mediated learning. Email: dan.frost@univ-savoie.fr

Alice Henderson, PhD, has been teaching English at the Université de Savoie in France since 1994. She strives to maintain a close link between her research and her teaching, which includes pronunciation classes for first year English majors, oral presentation skills for Psychology majors, and intensive courses for researchers who want to improve the intelligibility of their

spoken English for international conferences. She has also been fortunate enough to participate in teacher training programmes in Norway, Poland, Spain, Switzerland and the United States. Email: alice.henderson@univ-savoie.fr

Alexander Kautzsch, PhD, is a lecturer in English Linguistics in the Department of English and American Studies at the University of Regensburg, Germany. He teaches all kinds of linguistics classes (introductory classes, phonemic transcription, syntax, etc.) as well as practical language classes (esp. translation and grammar). One of his current research projects is concerned with the acquisition of L2 phonology by German learners of English. Email: Alexander.kautzsch@sprachlit.uni-regensburg.de

David Levey lectures at the University of Cádiz (Spain) on Phonetics and Phonology and Sociolinguistics. His main research interests include accents and language variation, bilingualism and second language acquisition. In addition to publications on various aspects of language teaching, he has written extensively on the sociolinguistics and language situation of Gibraltar. Email: david.levey@uca.es

Deirdre Murphy completed her Ph.D. in Applied Linguistics in Trinity College Dublin in 2011, on the interface between learner identity, motivation and autonomy in EFL pronunciation learning. Her current research interests include multilingualism and English language teaching and learning, with a particular focus on pronunciation and vocabulary, and she is a part of LUCIDE, a network of researchers investigating multilingual urban communities. She currently teaches English for Academic Purposes in Trinity College Dublin. Email: murphyd3@tcd.ie

Ewa Waniek-Klimczak is Professor of English Linguistics and Head of the Department of English Language and Applied Linguistics at the University of Łódź. Her main areas of interest are second language phonetics and phonology, sociolinguistics and pronunciation teaching. She organizes annual conferences on native and non-native Accents of English every December. She has edited and co-edited collections of papers on applied phonetics, with the most recent publication co-edited with Linda Shockey on *Teaching and researching English accents in native and non-native speakers* (2012, Berlin Heidelberg: Springer Verlag). Her main previous publications include the book on *Temporal parameters in second language speech: An applied linguistic phonetics approach.* (2005, Łódź: Wydawnictwo Uniwersytetu Łódzkiego) and *Issues in Accents of English I and II* (2008 and 2010, Cambridge Scholar Publishing). She is an editor-in-chief of *Research in Language*, an international journal published by the University of Łódź. Email: ewaklim@uni.lodz.pl

REFERENCES

Bradford, B. & Kenworthy. J. (1991). Phonology on teacher training courses. *Speak Out!, 9*, 12–15.

Breitkreutz, J. A., Derwing, T. M. & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal,* 19 (1), 51–61.

Brown, H. D. (2003). *Language Assessment: Principles and Classroom Practices*. London: Pearson Longman ESL.

Brown, A. & Taylor, L. (2006). A worldwide survey of examiners' views and experience of the revised IELTS Speaking Test. *Research Notes*, 26, 14–18.

Burgess, J. & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System,* 28 (2), 191–215.

Celce-Murcia, M., Brinton, D. M. and Goodwin J. M. (1996). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages.* Cambridge: Cambridge University Press.

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.

DeVelle, S. (2008). The revised IELTS Pronunciation scale. *Research Notes*, 34, 36–38.

Dimitrova, S. & Chernogorova, Ts. (2012). English pronunciation models and tertiary-level students: A Bulgarian perspective. In T. Paunović & B. Čubrović (Eds.) *Exploring English Phonetics* (pp. 208–224). Cambridge: Cambridge Scholars Publishing.

Dörnyei, Z. & Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing, 2nd Edition*. New York: Routledge.

Foote, J. A., Holtby, A. K. & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada. *TESL Canada Journal,* 29 (1), 1–22.

Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly*, 1 (4), 253–266.

Gass, S. & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of non-native speech. *Language Learning*, 34 (1), 65–89.

Henderson, A., Frost, D., Kautzsch, A., Kirkova-Naskova, A., Levey, D., Tergujeff, E. & Waniek-Klimczak, E. (2011). Joint session on English Pronunciation Teaching in Europe Survey. *Presentation given at the V International conference on native and non-native accents of English, December 17th, 2011, Łódź, Poland.*

Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Cunningham, U. & Curnick, L. (2012). The English Pronunciation Teaching in Europe Survey: Selected results. *Research in Language,* 10.1, 5–27.

Henderson, A., Kautzsch, A., Kirkova-Naskova, A., Frost, D., Levey, D., Murphy, D., Tergujeff, E., Curnick, L. & Waniek-Klimczak, E. (2013a). English pronunciation teaching in Europe: Inside & outside the classroom. *Presentation given at the Third International Conference on English Pronunciation: Issues and Practices (EPIP3), May 9th, 2013, Murcia, Spain.*

Henderson, A., Tergujeff, E., Kirkova-Naskova, A., Kautzsch, A., Waniek-Klimczak, E., Levey, D., Frost, D. & Curnick, L. (2013b). English Pronunciation Teaching in Europe Survey: Factors inside and outside the classroom. In Mompeán, J. & Fouz, J. (eds.), *Proceedings of the Third International Conference on English Pronunciation: Issues and Practices (EPIP3), Murcia 8th – 10th May 2013* (pp. 36–40). Murcia, Spain: University of Murcia.

Hubard, C., Gilbert, S., & Pidcock, J. (2006). Assessment processes in Speaking tests: a pilot verbal protocol study. *Research Notes*, 24, 14–19.

Huhta, A. (2010). Diagnostic and formative assessment. In Spolsky, B. & F. M. Hult (Eds.) *The Handbook of Educational Linguistics* (pp. 469–482). Oxford: Blackwell.

Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: OUP.

Kormos, J. & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145–164.

Levis, J. M. (2006). Pronunciation and the assessment of spoken language. In R. Hughes (Ed.), *Spoken English, TESOL and Applied Linguistics: Challenges for Theory and Practice* (pp. 245–270). New York: Palgrave MacMillan.

Macdonald, S. (2002). Pronunciation – views and practices of reluctant teachers. *Prospect: An Australian Journal of TESOL,* 17 (3), 3–18.

Milanović, M. (2002). *Common European framework of reference for languages: Learning, teaching, assessment: Language examining and test development*. Strasbourg, France: Council of Europe Language Policy Division.

Munro, M. & Derwing, T. (1999). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning*, 49, supp 1, 285-310.

Murphy, J. M. (1997). Phonology courses offered by MATESOL programs in the U.S. *TESOL Quarterly,* 31 (4), 741–764.

Murphy, D. (2011). An investigation of English pronunciation teaching in Ireland. *English Today,* 27 (4), 10–18.

Nowacka, M. (2010). The ultimate attainment of English pronunciation by Polish college students: A longitudinal study. In E. Waniek-Klimczak (Ed.) *Issues in Accents of English 2* (pp. 233–260). Newcastle upon Tyne: Cambridge Scholars Publishing.

Paunović, T. (2009). Plus ça change... Serbian EFL students' attitudes towards varieties of English. *Poznań Studies in Contemporary Linguistics*, 45 (4), 511–533. Retrieved from: http://versita.metapress.com/content/6563h54842u858nm/fulltext.pdf

Szpyra-Kozłowska, J., Frankiewicz, J., Nowacka, M. & Stadnicka, L. (2005). Assessing assessment methods - on the reliability of pronunciation tests in EFL. In *Online Proceedings from the Phonetics Teaching and Learning Conference (PTLC) 2005*. Retrieved from: http://www.ucl.ac.uk/psychlangsci/ptlc/proceedings_2005/ptlcp37

Tergujeff, E. (2012). English pronunciation teaching: Four case studies from Finland. *Journal of Language Teaching and Research,* 3 (4), 599–607.

Walker, R. (2010). *Teaching the Pronunciation of English as a Lingua Franca.* Oxford: OUP.

Walker, R. (1999). Proclaimed and perceived wants and needs among Spanish teachers of English. *Speak Out!,* 24, 25–32.

Waniek-Klimczak, E. (2002). Context for teaching English phonetics and phonology. In E. Waniek-Klimczak & P. J. Melia (Eds). *Accents and Speech in Teaching English Phonetics and Phonology* (pp. 139–152). Frankfurt am Main: Peter Lang.

Waniek-Klimczak, E. & Klimczak, K. (2005). Target in speech development: learners' views. In K. Dziubalska-Kołaczyk & J. Przedlacka (Eds.) *English Pronunciation Models: A Changing Scene* (pp. 229–250). Bern: Peter Lang.

Weir, C. J. (2005). Limitations of the Common European framework for developing comparable examinations and tests. *Language Testing*, 22 (3), 281–300.

Yates, L., Zielinski, B. & Pryor, E. (2011). The assessment of pronunciation and the new IELTS Pronunciation scale. *IELTS Research Reports*, 2, 1–46. [published online] Retrieved from: http://www.ielts.org/pdf/Vol12_Report1.pdf

ACOUSTIC CUES FOR ENGLISH LEXICAL STRESS PERCEPTION BY MANDARIN NATIVE SPEAKERS: A CRITICAL REVIEW

Sibo Chen, Simon Fraser University

Previous studies have suggested that Mandarin and English speakers use different phonetic cues in speech perception. In particular, there are many differences between English and Mandarin in terms of lexical stress patterns. Different tones are used to change lexical meanings in Mandarin, whereas such phenomena do not exist in English, which indicates potential inter-language interferences in the perception of English lexical stress by Mandarin English-as-an-Additional-Language (EAL) learners. The current paper discusses Mandarin EAL learners' perception of English lexical stress by critically reviewing previous studies. In general, findings have suggested that Mandarin EAL learners use various phonetic cues in English lexical stress perception, though which phonetic cues are primarily used is still under debate. The paper further identifies several variables that may account for the contrastive findings in previous studies and discusses their potential implications for EAL pedagogy development.

INTRODUCTION

Although the perception of English lexical stress by Mandarin EAL learners have been investigated in experimental settings (Archibald, 1997; Lai & Sereno, 2007; Ou 2004; Wang, 2008), previous studies have shown conflicting findings of the acoustic cues used by Mandarin EAL learners for English lexical stress perception. Although many studies have examined different teaching models in improving students' listening accuracy (Lai 2008; Abe, 2010), most of them have used English native speakers' judgment as the primary criterion to assess participants' improvement, which lacks accurate acoustic measurements.

Exploring the perception process of English lexical stress by EAL learners has significant theoretical and pedagogical implications. Theoretically, studying the acquisition process of lexical stress can reveal how the L2 acquisition process is constrained by interlanguage grammar. As suggested by Flege's (1995) Speech Learning Model (SLM), the mature L1 phonological system plays an active in the process of L2 phonological acquisition: similar sounds in the L1 and L2 may interfere with each other and create difficulties for learners. However, models like the SLM focus mainly on the segmental level of L2 phonological acquisition. Suprasegmental features such as lexical stress are not sufficiently specified. For example, although several studies have documented the phonological difficulties met by Mandarin EAL learners regarding Mandarin-English phonological interactions (Yao, 2008; Wang, 2008; Huang & Radant; 2009), it is only in recent years that EAL lexical perception has attracted increasing attention from L2 phonetics researchers (Lai & Sereno, 2007; Wang, 2008; Yu & Andruski, 2010; Zhang, 2008).

Pedagogically, the acquisition of English lexical stress is very important in EAL instruction since it is closely associated with learners' listening comprehension accuracy. Misperception of stressed syllables in English may lead to inaccurate comprehension, thereby resulting in ineffective communication (Liu, 2007). However, relevant discussions of lexical stress are rarely found in EAL instruction texts and little effort has been made to transfer relevant experimental findings to language classrooms (e.g. Doughty & Long, 2003; Levis, 1999; Lightbown & Spada, 1999; Nuan, 1999).

This paper contributes to these issues by critically reviewing recent studies on the perception of English lexical stress by Mandarin EAL speakers. The review has three major purposes. First, it explores factors that contribute to the marginal status of phonological accuracy studies in applied linguistics, aiming at identifying current gaps on English lexical stress acquisition between classroom-based and experimental studies. Second, it reviews experimental studies on these issues and evaluates their findings, especially on the acoustic cues used by Mandarin EAL speakers for English lexical stress perception. Third, it discusses the potential of transferring experimental findings to real language teaching contexts.

The paper first reviews research on lexical stress perception from an applied linguistics perspective, then compares Mandarin and English lexical stress systems, then examines recent empirical explorations on English lexical stress perception by Mandarin native speakers, followed by the concluding sections.

The Marginalized Status of Phonological Accuracy Studies in Applied Linguistics

English lexical stress acquisition has been somewhat marginalized in previous language pedagogy studies despite its significant roles in EAL listening comprehension and pronunciation (Derwing & Munro, 2005). This phenomenon may partially be attributed to the prevalence of Communicative Language Teaching (CLT) in current EAL education (Doughty & Long, 2003; Lightbown & Spada, 1999; Nuan, 1999). One key factor of CLT was its emphasis of language content rather than form. As a result, lexical stress, along with other elements of language structure, was deemphasized in early CLT frameworks.

Lexical stress is an understudied topic in applied linguistics in general, and the learning of lexical stress for Mandarin learners has not been sufficiently studied. In the following sections, I will focus on experimental studies on English lexical stress perception by Mandarin native speakers and review issues concerning the status of acoustic cues in Mandarin-English phonological interference.

Mandarin and the Lexical Stress Systems of English

Lexical stress can be defined as "the syllable prominence in a word" (Ou, 2004, p. 1541). Generally, the stress systems of natural languages can be presented as the taxonomy in Figure 1, in which languages are categorized as stress, pitch accent, or tone languages (Altmann & Vogel, 2002; Archibald, 1997). "Pitch accent" refers to the use of pitch to mark syllabic prominence whereas "tone" refers to the use of pitch in language to distinguish lexical or grammatical

meanings. It has been argued that all stress languages belong to accentual languages and tone languages are all non-accentual languages. Based on this taxonomy, the primary difference between English and Mandarin in lexical stress is that English is an accentual language with movable stress assignment while Mandarin uses tone lexically to differentiate word meanings.
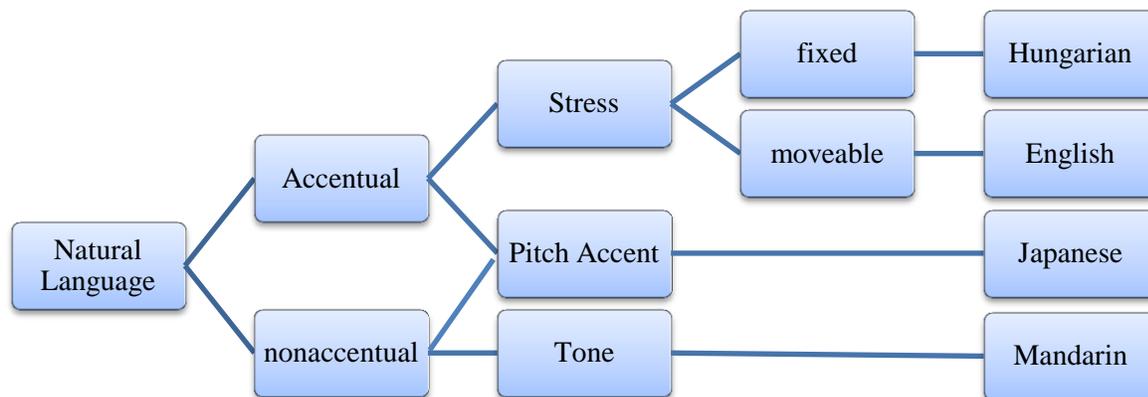
Figure 1. A taxonomy of stress systems in natural languages (adapted from Archibald, 1997)

The acoustic cues for English lexical stress have been extensively studied by researchers (e.g. Fry, 1955; 1958; Lehiste, 1970). Generally, four acoustic cues have been identified for English lexical stress: fundamental frequency (F0), duration, intensity and formant structure (Wang, 2008). Fundamental frequency (F0) is defined as the lowest frequency of a periodic waveform; duration is the time interval; intensity refers to the power of sound; and formant structure is the acoustic resonance of sounds (Hirst, 2006). In English lexical stress perception, the four cues are not weighted equally. Series of studies have shown that the F0 is the most influential factor, followed by duration and intensity. For instance, in the seminal study of Fry (1958), F0 has a consistent effect on duration manipulated word "subject." The research participants tended to judge the lexical stress as occurring word initially when F0 was higher in the first syllable. Fry (1955) also showed that when judging the stress of synthesized words with various duration and intensity manipulations, native speakers' performance was influenced by both duration and intensity cues. Lehiste (1970), reviewing Fry's studies, argued that the F0 was a very strong cue for lexical stress identification while the function of duration and intensity was secondary. A fourth acoustic cue, formant structure has also been discussed in recent studies. For example, it has been suggested that English native speakers have different preferences for the stress patterns in nouns and verbs based on formant structure (Lai & Sereno, 2007; Yu, 2008; 2010).

There are three theories regarding the Mandarin lexical stress system: the non-stress theory, final stress theory, and left-headed theory (Lin, 2001). The non-stress theory says that in Mandarin there is no lexical stress since tones are used lexically. The final stress theory argues that in

typical disyllabic Mandarin phrases[1], stress is word final, as indicated by the longer duration of the second syllable (Chao, 1968). The third theory is the left-headed theory which proposes that Mandarin has initial stress. This theory is supported by the longer onset time and wider pitch of the first syllable (Duanmmu, 2000) as well as other evidence such as the distribution of the neutral tone in Mandarin (Lin, 2001). Lin (2001), however, made three arguments against the final-stress theory: first, in disyllabic Mandarin phrases, the first syllables does not exhibit tone reduction; second, the lengthening of the final syllables may be attributed to a domain-final effect, and third, the lengthening of the final syllables are not observed in polysyllabic sentences (Wang & Wang, 1993). For the present study, I will adapt the left-headed theory.

Despite the various theories on Mandarin lexical stress system, previous research on Mandarin tones has shown that Mandarin native speakers use acoustic cues (F0, duration, and intensity) in Mandarin tone perception and production (e.g. Howie, 1976; Lin, 1988; Moore & Jongman, 1997; Tseng, 1990), which resembles English lexical stress identification. F0 is regarded as the primary acoustic cues for Mandarin tones. For instance, Howie (1976) used synthetic speech with manipulation on the F0 contour in three identification tests and found that participants achieved the best performance when the pitch patterns are maintained. Similar results were reported in Lin (1988), in which the highest discrimination rate in synthesized speech tests (tests using artificially created human speech) was provided by F0. By comparison, the other two acoustic cues (duration & intensity) are believed to have limited influence on Mandarin tone identification. In Lin's (1988) discrimination test, the influence of different duration of the four tones was estimated to contribute to only 3% of the results. Fu & Zeng (2000) suggested that the intensity contour in Mandarin tones is highly correlated with the F0 contour, which makes it the secondary significant factor in Mandarin tone perception.

In sum, acoustic cues (F0, duration, and intensity) are utilized differently in Mandarin and English. These are used for lexical stress identification in English while in Mandarin they are used lexically to differentiate word meanings. Thus, it is expected that for Mandarin EAL learners, their perception of English lexical stress will be influenced by their L1 experience with these acoustic cues. One possibility is that Mandarin EAL learners will transfer their acoustic use in Mandarin into English and thereby memorizing the English stress lexically (Archibald, 1997). Alternatively, they may acquire the use of acoustic cues in English and show similar lexical stress identification pattern as English native speakers.

Previous Studies on English Lexical Stress Perception by Mandarin EAL Speakers

The perception of English lexical stress has been studied among EAL learners with multiple L1 backgrounds, such as Spanish (Archibald, 1993; Guion, Harada, & Clark, 2004), Korean (Guion, 2005), and Mandarin (Archibald, 1997). In general, these studies have confirmed the influence of

---

[1] Many Chinese linguists (e.g. Chao, 1968) argue that the essential word unit in Chinese is disyllabic; such as the phase "Chi-Fan" that is the composition of "Chi" (eat) and "Fan" (rice). For the consistency of discussion, in the following paper, the "lexical stress" of Chinese refers to the stress patterns in typical disyllabic Chinese phrases.

L1 lexical stress systems on EAL learners' perception of English lexical stress. For instance, Archibald (1993) investigated the acquisition of English metrical parameters by adult Spanish speakers and found that the extrametricality markings[2] in Spanish are transferred in to the participants' L2 English. Similarly, in Guion (2005), both early and late Korean-English bilinguals demonstrated non-native like knowledge of the distributional patterns of stress placement across the lexical classes of noun and verb in English.

One seminal early work on L2 lexical stress acquisition is Dresher and Kaye (1990), which adopted the Principles and Parameters framework by Chomsky and proposed that there were eight parameters on language's phonological system (Table 1). According to this model, different languages employ different values on these parameters, which generate various lexical stress systems. Thus for a L2 learner, acquisition of L2 lexical stress patterns means adjusting the L1 parameters to the L2 settings.

Table 1

Parameters of Phonological Stress System

P1	The word-tree is strong on the [left/right].
P2	Feet are [binary/unbounded].
P3	Feet are built from the [left/right].
P4	Feet are strong on the [left/right].
P5	Feet are quantity-sensitive [yes/no].
P6	Feet are quantity-sensitive to the [rime/nucleus].
P7	A strong branch of a foot must itself branch [yes/no].
P8	There is an extrametrical syllable [yes/no].
P8A	It is extrametrical on the [left/right].

Adopted from Dresher and Kaye (1990)

Following Dresher & Kaye (1990), Archibald (1997) applied this model to investigate the English lexical stress perception and production by Mandarin native speakers in a longitudinal study. He found that in both perception and production tests, the Mandarin subjects showed no acquisition of English stress assignment principles, which indicated that the Mandarin subjects tended to treat English lexical stress as a phenomenon that requires mechanical memorization. Archibald further argued that this could be explained by the different linguistic processing of pitch in Mandarin and English. The fact that pitch is phonemic in Mandarin may transfer to the Mandarin EAL learners' perception of English. This is in accord with studies of lexical stress in other language interactions, such as French speakers who learn Spanish as an L2. In a series of studies of Spanish stress perception by French speakers, Dupoux et al showed that native speakers of French exhibited stress 'deafness.' They had difficulties distinguishing stress contrasts in Spanish, which can be partially attribute to the fact that stress in French is

---

[2] Extrametricality refers to the phenomena that in certain languages, the rightmost syllable of words can be 'ignored' when metrical structure is constructed (Hulst, 2006).

non-contrastive while it is contrastive in Spanish (Dupoux, Pallier, Sebastián, & Mehler, 1997; Dupoux, Peperkamp, & Sebastián-Gallés, 2001; Dupoux, Sebastián-Gallés, Navarrete, & Peperkamp, 2008).

By comparison, studies under different frameworks have shown conflicting results against the strong argument of Archibald (1997). These studies have demonstrated that at least to some extent, Mandarin EAL learners are able to use acoustic cues in English lexical stress perception. For instance, Ou (2004) conducted perception test of trisyllabic English pseudo-words with 20 postgraduate Mandarin students and found that eight showed sensitivity to extrametricality of different grammatical categories and syllable weight, which indicates some sort of metrical computation of English stress. This result was further tested in Ou (2010) with Taiwanese EAL learners and the results supported the hypothesis that Mandarin EAL learners may be able to improve their phonological awareness of the differences between lexical tone and lexical stress according to their developmental stages.

Furthermore, several recent experimental studies have explicitly investigated the use of acoustic cues by Mandarin EAL learners in English lexical perception and their results generally confirmed Mandarin EAL learners' ability to use acoustic cues in English lexical perception. Wang (2008) used pseudo-words with manipulation in F0, duration, and intensity to test Mandarin EAL learners and English native speakers. Results indicated that, similar to English native speakers, the Mandarin EAL learners showed systematic variation in concordance with the manipulation of acoustic cues. Other evidence is provided by Yu and Andruski (2010), in which real words and hums were also tested along with pseudo-words. The results demonstrated that Mandarin speakers had different response patterns to trochaic and iambic stress in the three types of stimuli, which indicated their sensitivity to stress structure in English.

Other studies have disagreed on the weight of different acoustic cues in Mandarin EAL learners' perception of English lexical stress. In Wang (2008), the Mandarin EAL participants had significantly lower reliance scores in duration and intensity, but higher for F0, which indicated there was a transfer of reliance of F0 in Mandarin tone identification to English lexical stress perception. In other words, the results supported the hypothesis that Mandarin EAL learners can actively redeploy their knowledge of acoustic cues in English stress perception. This result is also supported by Lai (2008), in which advanced Mandarin EAL listeners are found to focus more on maximum F0. In contrast, in Zhang (2007), both Mandarin EAL learners' perception and production of English stress were tested and Mandarin EAL learners used F0 and intensity to a lesser extent while vowel quality was weighted greater than other acoustic cues. Finally, formant structure was also found to influence Mandarin EAL learners' preference of stress perception. In a series of studies, EAL learners identified English pseudo-words with iambic stress patterns significantly more quickly than those with trochaic patterns, which supports the transfer of Mandarin word-final stress to EAL phonological perception (Yu, 2008; Yu & Andruski, 2010; Yu & Sereno, 2007).

The recent experimental studies have provided a complicated picture regarding the acoustic cues used by Mandarin EAL learners in English lexical stress perception. Although all the major

acoustic cues (i.e., F0, duration, intensity, vowel quality, and formant structure) contributed to the perception of English stress, which acoustic cue played the primary role is still under debate.

There are several possible reasons for the varied results. First, the experimental subjects' proficiency of English may have influenced their performance in the perception tests. Although all the studies used advanced Mandarin EAL learners as participants, their exposure to English varied from study to study. In Wang (2008), the participants were college students majoring in English at a Mandarin university; thus their exposure to English was limited to classroom instruction. By comparison, Zhang (2007) and Yu and Andruski (2010) used Mandarin speakers studying in US universities. Their residence in an ESL context may have improved their awareness of English lexical stress. As shown in the comparative analysis in Lai (2008), beginning Mandarin EAL learners tended to rely on duration to identify English lexical stress, which was not found for the advanced EAL group. Ou (2010) also provided evidence to support the claim that phonological training can improve Mandarin EAL learners' phonological awareness. Second, the use of pseudo-words may be a factor since using real words may introduce semantic processing. In Archibald (1997), the test tokens were real words such as "aroma", and "cinema", and one possible interpretation of his results might be that the participants' familiarity with some of the tokens made them process the tokens phonemically. In more recent studies (e.g., Ou, 2010; Wang, 2008), pseudo-words were used as the stimuli. Yu and Andruski (2010) showed that participants' performance did not vary significant between pseudo and real words, which suggested that the use of pseudo-words does not comprise experiment subjects' performance and thus may be regarded as a more reliable measurement than using real word stimuli. Finally, various interpretation of the Mandarin lexical stress system may also lead to conflict experimental findings. As reviewed in the previous section, there are three contrastive theories of the Mandarin lexical stress system. If we adapt the non-stress theory, then results indicating no transfer between Mandarin and English stress are predicted. By contrast, if the existence of lexical stress is assumed in Mandarin, then we would expect some degree of transfer since Mandarin EAL learners are not "stress deaf" under this basis. In the reviewed studies, Archibald (1997) was based on the non-stress theory while Ou (2004) and Wang (2008) adapted the left-headed one. In the series studies by Yu, the final stress theory is used. Such distinctive starting points may affect the follow-up data interpretation. To sum up, the participants' proficiency of English, the stimulus type and the assumptions about Mandarin lexical stress system may also have led to conflicting findings in previous studies. Thus in future studies, these factors should be thoroughly considered in advance.

Many empirical questions regarding Mandarin speakers' perception of English lexical stress are still unsolved. For instance, although experimental studies have shown Mandarin speakers are able to use acoustic cues in processing English lexical stress, the process used is still unclear. One possible situation may be that the Mandarin speakers redeploy their knowledge of acoustic cues in L2 lexical stress; but it is also possible that they are treating stress in English lexicons like tones. Finally, the developmental stages of Mandarin EAL speakers' lexical stress system have not been described.

PEDAGOGICAL IMPLICATIONS AND CONCLUSION

Although pedagogical implications are not the focus of the reviewed studies, their findings provide valuable empirical evidence for EAL's listening comprehension instruction. First of all, all the studies address the importance of raising Mandarin EAL learners' awareness of suprasegmental features in English. Furthermore, the use of pseudo-words in perception tests may also be modified for language assessment. To avoid the familiarity effect, less common words should be used when testing EAL learners' perception of English stress in classroom settings.

In sum, previous studies on Mandarin EAL learners' perception of English lexical stress suggested complicated patterns of acoustic cue-use during identification. Overall, most previous studies suggested that Mandarin EAL learners were able to actively use acoustic cues in processing English lexical stress. However, the studies contradict each other in the weight of different acoustic cues, which indicates a need for careful control of participants' English proficiency, appropriate construction of stimuli, and careful selection of Mandarin lexical stress theory.

ACKNOWLEDGEMENT

ABOUT THE AUTHOR

Sibo Chen is a graduate student at the School of Communication, Simon Fraser University, Canada. He received his MA degree in Applied Linguistics from the Department of Linguistics, University of Victoria, Canada. His major research interests include corpus linguistics, discourse analysis, and the interactions between language and communication.

REFERENCES

Abe, H. (2010). Form-focused instruction in L2 pronunciation pedagogy: The effect of negotiation of form in a Japanese classroom. *Proceeding of New Sound 2010.*

Altmann, H. & Vogel, I. (2002). L2 acquisition of stress: the role of L1. Paper presented at the DGfS Annual Meeting "Multilingualism Today" in Mannheim, Germany, March 2002.

Archibald, J. (1993). Metrical phonology and the acquisition of L2 stress. In F. R. Eckman (Ed.), *Confluence: Linguistics, L2 acquisition and speech pathology* (pp. 37-48). Amsterdam, NL: John Benjamins.

Archibald, J. (1997). The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics, 35*(1), 167-181.

Chao, Y. R. (1968). *A spoken grammar of Mandarin*. Berkeley, CA: University of California Press.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-379.

Doughty, C. J., & Long, M. H. (Eds.). (2003). The handbook of second language acquisition. Malden, MA: Blackwell.

Dresher, B. E., & Kaye, J. D. (1990). A computational learning model for metrical phonology. *Cognition, 34*(2), 137-195.

Duanmu S. (2000). *The phonology of standard Mandarin*. New York, NY: Oxford University Press.

Dupoux, E., Pallier, C., Sebastián-Gallés, N. & Mehler, J. (1997). A destressing 'deafness' in French? *Journal of Memory and Language, 36*, 406-421.

Dupoux, E., Peperkamp, S., & Sebastián-Gallés, N. (2001). A robust method to study stress "deafness". *The Journal of the Acoustical Society of America, 110*(3), 1606-1618.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition, 106*(2), 682-706.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399-423.

Flege, J. E. (1995). Second language speech learning theory, findings and problems. In W. Strange (Ed.) *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (pp. 233-277). Baltimore, MD: York Press.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765-768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126-152.

Guion, S. G. (2005). Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition, 27*(4), 503-533.

Guion, S. G., Harada, T., & Clark, J. J. (2004). Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition*, *7*, 207–226.

Hirst, D. J. (2006). Prosodic aspects of speech and language. In K. Brown (ed.), *Encyclopedia of Language and Linguistics* 2nd Ed. (pp. 167-178), Oxford, UK: Elsevier.

Howie, J. M. (1976). *Acoustical Studies of Mandarin Vowels and Tones*. Cambridge, UK: Cambridge University Press.

Huang, H., & Radant, J. (2009). Mandarin phonotactic patterns and the pronunciation difficulties of mandarin-speaking EFL learners. *Asian EFL Journal, 11*(4), 148.

Hulst, H. V. D. (2006) Word stress. In K. Brown (ed.), *Encyclopedia of Language and Linguistics* 2nd Ed. (pp. 655-665), Oxford, UK: Elsevier.

Jiang, H. (2010). Effect of L2 phonetic learning on the production of L1 vowels: a study of Mandarin-English bilinguals in Canada. *Proceeding of New Sound 2010*.

Lai, Y. (2008). *Acoustic realization and perception of English lexical stress by mandarin learners.* (Ph.D., University of Kansas). *ProQuest Dissertations and Theses*.

Lai, Y., & Sereno, J. (2007). Acquisition of lexical stress by Mandarin learners. *The Journal of the Acoustical Society of America, 121*(5), 3071.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: M.I.T. Press.

Levis, J. (1999). Intonation in theory and in practice, revisited. *TESOL Quarterly, 33*, 37-54.

Lightbown, P., & Spada, N. (1999). *How languages are learned (Rev. ed.)*. Oxford, England: Oxford University Press.

Lin, Hua (2001). Stress and the distribution of the neutral tone in Mandarin. In D. B. Xu (ed.). *Mandarin Phonology in Generative Grammar* (pp. 139-161). New York, NY: Academic Press.

Lin, M. C. (1988). Putong hua sheng diao de sheng xue texing he zhi jue zhengzhao [Standard Mandarin tone characteristics and percepts]. *Zhongguo Yuwen*, *3*, 182-193.

Liu, L. (2007). *The English deaf: An orthoepical study of the effectiveness of stress placement intervention for Mandarin English speakers*. Dissertation Abstracts International, A: The Humanities and Social Sciences.

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of mandarin Mandarin tones. *Journal of the Acoustic Society of America, 102*, 1864-1877.

Ou, S-C. (2004). An optimality-theoretic approach to word stress: evidence from Mandarin-English interlanguage. Proceeding of 15th ICPhS at Barcelona.

Ou, S-C (2010). Taiwanese EFL Learners' Perception of English Word Stress. *Concentric: Studies in Linguistics, 36*(1), 1-23.

Tseng, Chiu-yu. (1990). *An Acoustic Phonetic Study on Tones in Mandarin Mandarin*. Taipei: Institute of History & Philology Academia Sinica.

Wang, C. (2004). Self-concept, English pronunciation and EFL learning. *Foreign Language Teaching and Research, 36*(1), 56-63.

Wang, J. & Wang, L. J. (1993). Putonghua duoyinjie ci yinjie shichang fenbu moshi [The types of relative length of syllables in polysyllabic words in Putonghua]. *Zhonghuo Yuwen [Mandarin Philology]. 1993*(2): 112-116.

Wang, Q. (2008). *Perception of English stress by mandarin Mandarin learners of English: An acoustic study*. Dissertation Abstracts International, A: The Humanities and Social Sciences.

Yao, H. (2008). *An interview study of native Mandarin-speaking English as a foreign language university students' experience of acquiring English pronunciation.* Dissertation Abstracts International, A: The Humanities and Social Sciences.

Yu, V. Y., & Andruski, J. E. (2010). A cross-language study of perception of lexical stress in English. *Journal of Psycholinguistic Research, 39*(4), 323-344.

Yu, V. Y., & Andruski, J. E. (2011). The effect of language experience on perception of stress typicality in English nouns and verbs. *The Mental Lexicon, 6*(2), 275-301.

Yu, Y. (2008). *A cross-language study in perception of lexical stress in English.* (Ph.D., Wayne State University). ProQuest Dissertations and Theses.

Zhang, Y. (2008). *Production and perception of English lexical stress by native mandarin speakers.* (Ph.D., Purdue University). ProQuest Dissertations and Theses.

Zhang, Y., & Francis, A. (2010). The weighting of vowel quality in native and non-native listeners' perception of English lexical stress. *Journal of Phonetics, 38*(2), 260.

CONFUSION AS A COMPLEMENT TO INTELLIGIBILITY RESEARCH

Ettien Koffi, St. Cloud State University

Confusion research has been used for more than 50 years to test speech intelligibility in automatic speech recognition systems. In this paper, I apply its methodology and its findings to L2 English intelligibility research. Preliminary findings indicate that the Perceptual Distance Hypothesis (PDH) can help to predict vocalic substitutions that impinge on intelligibility and those that do not. Furthermore, PDH confirms findings by Derwing, Flege, Munro, and others, that the English vowels [æ, ʌ, ɑ] produced by many nonnative talkers are more likely to impinge on intelligibility than other vowels. I conclude, hesitantly and pending further research, that the available data confirms that confusion findings and intelligibility findings complement each other.

INTRODUCTION

The concept of intelligibility that Derwing and Munro have championed for about two decades has brought both enthusiasm and scrutiny among L2 pronunciation teachers and researchers. Levis (2005, p. 370) notes that the "Intelligibility Principle" has begun to assert itself as a dominant force to be reckoned with. Dauer (2005, p. 548) agrees with this assessment, yet she contends that intelligibility is "difficult to define and measure." Indeed, Levis (2011)   reports that a panel of experts did not come to a consensus on the factors that contributed to the unintelligibility of two L2 speech samples. It is, therefore, not an accident that the main theme of the 2012 PSLLT conference deals with ways of assessing intelligibility reliably. In this paper, I review the methodology employed in intelligibility research and suggest that confusion research can be used in tandem with intelligibility research to assess the latter more reliably. Confusion research has contributed significantly to both speech science and speech recognition research. However, it is only now that Miller and Nicely's (1955) groundbreaking work is trickling down into mainstream phonetic and phonology textbooks. Ladefoged and Disner (2012, pp. 99-113) and Johnson (2012, pp. 112-127) have included a chapter on confusion in the latest editions of their popular phonetic textbooks. I contend in this paper that confusion research can complement intelligibility research because it provides principled answers to many of the issues that intelligibility researchers are confronted with. For instance, in standard intelligibility methodology, native speaker judges are asked to render judgments as to whether or not a non-native production of a specific word is intelligible or not. If the word is deemed unintelligible, the researcher catalogues all the unintelligible segments and looks for a pedagogical solution to the problem. However, confusion researchers are also interested in identifying unintelligible sounds, but their quest does not stop there. They probe further to find out which other sounds in English the unintelligible sound is confused with. Their findings help establish confusion rates. They also help predict the likelihood of confusion between a given phone and all other phones in the language.

Intelligibility Research: Its Goals, Scope, and Blind Spots

Derwing and Munro's research agenda centers around a lexical trio: Intelligibility, comprehensibility, and accentedness. On paper, the distinction between these three concepts

seems straightforward. However, during active listening, it is quite difficult to tease out the causality relationship between them, as noted by Jun and Li (2010, p. 58):

> The more interesting finding seems to be that when the raters were rating for accentedness, they mentioned comprehensibility, and in the same way, when the raters were rating for comprehensibility, they mentioned accentedness. This could point to the possible interrelationship, or "quasi-dependency" between comprehensibility and accentedness.

This echoes what Dauer (2005, p. 548) wrote almost a decade ago, namely that, "Although everyone states that the goal of pronunciation is intelligibility, it is difficult to define and to measure." It is indeed difficult to diagnose, let alone predict, how, when, and why an L2 pronunciation may be accented, incomprehensible, or unintelligible. During the 2010 PSLLT conference, Levis, the conference organizer, put together a panel of nine intelligibility researchers and asked them to listen and assess the speech of two international students. The panelists found it extremely hard to pinpoint what exactly caused intelligibility problems. Levis (2011, pp. 56-57) summarized the results of the panel discussions as follows:

> All participants listened to the free speech followed by the read speech of each student. None of the listeners had heard the recordings before. Everyone was provided with an unmarked copy of the reading passage for each student and with the questions the students were asked in the interview. Each recording was played once (for time's sake), and the panelists and audience members were asked to identify the features of the speech that most influenced the speaker's intelligibility or lack thereof. The task was quite difficult (even for the experts). Listening once was not felt to be enough, and everyone thought that another opportunity to listen would have been helpful.

It is actually good that the experts did not have another opportunity to listen to the speech sample, because in real life, more often than not, we have only one chance to listen to a speaker and make sense of what he/she is saying. For fear of appearing rude or causing L2 talkers to lose face about their language abilities, L1 hearers refrain from asking them to keep repeating themselves even if they do not understand a fair amount of what is being said. The very fact that the experts were not able to offer an accurate diagnosis upon hearing the speech sample only once speaks volumes about the fact that it is easier to describe intelligibility academically than to use it as a yardstick to gauge and predict how and why an L2 speaker may be unintelligible to native speakers of General American English (GAE). One can sense a tinge of disappointment in Levis' report as he wrote:

> They [the experts] were supposed to be able to judge which pieces of the student's performance were causing the biggest problems, and they found themselves not able to get a handle on the speech patterns. (Levis, 2011, p. 57).

Two pages later, Levis writes, "What was clear was that even the experts do not completely agree on what most impacted intelligibility." I contend in this paper that insights from confusion research can help assess intelligibility more effectively.

Goals and Scope Confusion Research

In 1955 Miller and Nicely published an article entitled "An analysis of perceptual confusion among some English consonants." Since the publication of this article, confusion research has emerged as a fruitful area of scholarship in automatic speech recognition (ASR) and in speech

perception studies. Confusion research depends crucially on the notion of perceptual distance between segments. This has led to the postulation of the Perceptual Distance Hypothesis (PDH), which I summarize and formulate for GAE as follows:

Perceptual Distance Hypothesis

> GAE segments that are in close perceptual proximity, *and* have a marginal relative functional load may be confused at no detriment to intelligibility, except as carriers of accentedness.

This postulate can be very useful in assessing the intelligibility of L2 English. In applying this concept, it is worth remembering that the perception of vowels is trickier than that of consonants. Ferrand (2007, pp. 263, 265, 287) suggests that this may be due to the fact that formant frequencies are the only cues that hearers have to distinguish between vowels. Johnson (2012, p. 108) describes an experiment in which the auditory perception of vowels was shown to be gradient, that is, the boundary lines between vowels often overlapped. Since vowels are perceptually harder to account for, if confusion findings can deal satisfactorily with vowels, then explaining the unintelligibility of consonants will be relatively easier. This is the reason why this paper is devoted exclusively to vowels.

Peterson and Barney (1952, p. 182) and Hillenbrand, Getty, Clark, and Wheeler (1995, p. 3108) provide the following confusion data about GAE vowels. Let's use the vowel [ɪ] to illustrate how the information in Tables 2 and 3 is interpreted by confusion researchers.   Many GAE speakers produced the vowel [ɪ] (Spoken Stimuli). Many other GAE hearers were asked to indicate which vowel they heard (Perceived Stimuli). Overall, 92.8% of the hearers perceived [ɪ] accurately, but 0.5% of the hearers perceived it incorrectly as [ɪ], another group of hearers misperceived it as [ɛ] 6.7% of the time, some confused it with [æ] 0.1% of the time, and other hearers thought they heard [ɚ] 0.25% of the time.

Table 2
Peterson and Barney's Vowel Confusion Matrix

		Perceived Stimuli									
		[i]	[ɪ]	[ɛ]	[æ]	[ɑ]	[ɔ]	[ʊ]	[u]	[ʌ]	[ɚ]
	[i]	99.9	.03	.05			.02				
	[ɪ]	.05	92.8	6.7	.01	>.01	>.01				.25
	[ɛ]		2.5	87.7	9.2	>.01	.02			.01	.49
Spoken Stimuli	[æ]		>.01	2.9	96.5	.01	.01			.14	.37
	[ɑ]		>.01		.18	87.0	9.8	.67		2.2	.06
	[ɔ]			>.01	.01	5.7	92.8	.69	.04	.60	.13
	[ʊ]			>.01	>.01	.15	.49	96.5	.93	1.6	.18
	[u]			>.01		.01		.75	99.1		.01
	[ʌ]		>.01	>.01	.07	5.2	1.2	1		92.2	.2
	[ɚ]			.22	.05	.01	.02			.01	99.6

Note: Adapted from Peterson and Barney (1952, p. 182).

Table 3

Hillenbrand et al. (1995) Vowel Confusion Matrix

		[i]	[ɪ]	[e]	[ɛ]	[æ]	[ɑ]	[ɔ]	[o]	[ʊ]	[u]	[ʌ]	[ɚ]	Total
							Perceived stimuli							
	[i]	99.6	0.1	0.1									0.1	99.9
	[ɪ]		98.8	0.2	0.9									99.9
	[e]	0.6	0.3	98.3	0.3		0.2	0.1		0.1				99.9
	[ɛ]		0.5		95.1	3.7	0.2	0.1		0.1		0.2	0.1	100
	[æ]				5.6	94.1	0.2							99.9
Spoken Stimuli	[ɑ]		0.1		0.1	0.3	92.3	3.5	0.1	0.2	0.1	3.3		100
	[ɔ]						13.8	82.0		0.1	0.1	3.8	0.2	100
	[o]							0.1	99.2	0.5			0.2	99.5
	[ʊ]				0.1					97.5	1.3	1.0		99.8
	[u]								0.4	1.9	97.2		0.1	99.6
	[ʌ]						3.7	1.8	0.3	3.2	0.2	90.8	0.2	100.2
	[ɚ]	0.1				0.2				0.2			99.5	100

Note: Adapted from Hillenbrand et al. (1995, p. 3108).

On the basis of the information in Tables 2 and 3, it is possible to group GAE vowels into three perceptual categories: Perceptually close vowels, perceptually overlapping vowels, and perceptually distant vowels. This three-way division provides the basis for the discussions in the remaining sections of the paper.

Perceptually Close Vowels
The vowels [ɑ] and [ɔ] are perceptually very close. When these two vowels are produced, GAE hearers have a hard time telling them apart from each other. In Peterson and Barney (1952), the vowel [ɑ] was accurately perceived as [ɑ] 87% of the time. The vowel [ɔ] was also perceived accurately 92.8% of the time. The rest of the time, they were confused with each other. The vowel [ɑ] was confused with [ɔ] 9.8% of the time, while [ɔ] was confused with [ɑ] 5.7% of the time. The confusion rate of these two vowels is much higher in the Midwest according to Hillenbrand et al. (1995). In their data [ɔ] was confused with [ɑ] 13.8% of the time, whereas [ɑ] was confused with [ɔ] only 3.5% of the time. Many commentators, including Small (2005, pp. 74-75) have noted that the distinction between [ɔ] and [ɑ] is not made in the Midwest, the western United States, and in parts ofNew England. Consequently, pairs of words such as *cot/caught, tot/taught,* and *Don/Dawn* sound the same to the speakers of these dialects.

Formant frequency data from 22 Central Minnesota female speakers shows clearly that [ɑ] and [ɔ] have merged, as shown in Figure 1.
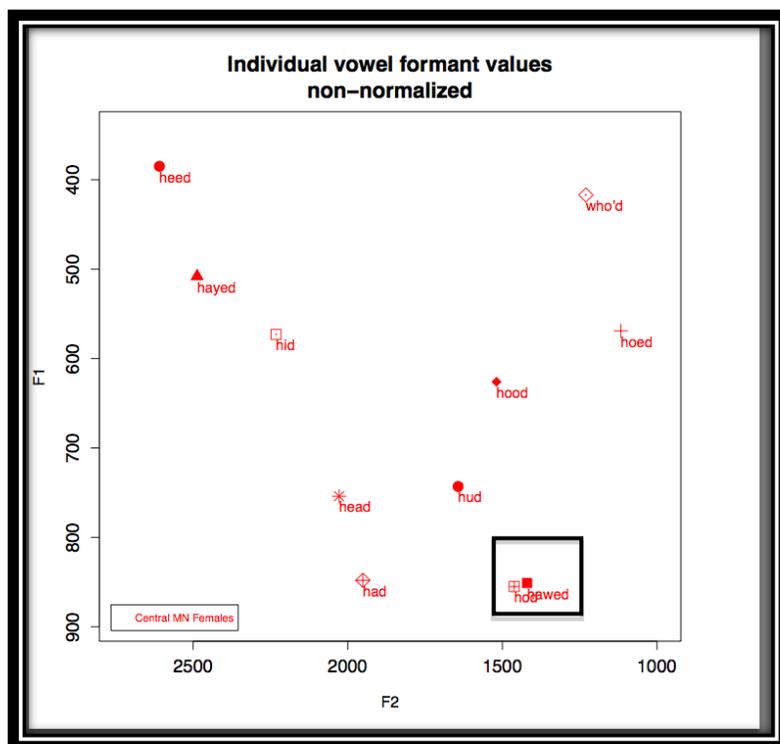
Figure 1. Acoustic Vowel Space of Central MN Females

The F1 formant of [ɑ] is 855 Hz, while that of [ɔ] is 851 Hz. The F2 formant of [ɑ] is 1420 Hz, and that of [ɔ] is 1462Hz. According to Baart (2010, p. 67), for two sounds to be perceived as acoustically different, there must be at least 200 Hz difference between the two F1s, and about 400 Hz difference between the two F2s. So, the difference of 4 Hz is insignificant for the two F1s, and so is the difference of 42 Hz for the two F2s. Therefore, acoustically speaking, [ɑ] and [ɔ] have merged in Central Minnesota. The merger of these two sounds was also reported in Ladefoged (2006, p. 89) who wrote that, "Many Midwestern and Californian speakers do not distinguish between [ɑ] and [ɔ]." Though Catford (1987, p. 90) rates the relative functional load of [ɑ] and [ɔ] at 26%, he also notes that in some dialects, the distinction is not made. Further evidence that [ɑ] and [ɔ] have merged in the speech of some people is found in Celce-Murcia, Brinton, Goodwin, & Griner (2010, p. 120). They instruct ESL teachers that /ɑ/ can be substituted for [ɔ] when teaching nonnative speakers. The Perceptual Distance Hypothesis proposed earlier works well for these two vowels; that is, since they have merged in many dialects of GAE, confusing one with the other, and vice versa, does not lead to unintelligibility.

Perceptually Distant Vowels

The confusion data found in Peterson and Barney (1952) and Hillenbrand et al. (1995) suggest that the features [± tense] and [± back] are used by GAE speakers to perceive perceptual distance. GAE hearers hardly ever confuse tense and lax vowels. Peterson and Barney (1952) did not study the vowels [e] and [o] in their data because they took them to be diphthongs. However, since in the Midwest, most people do not perceive these vowels as diphthongs, Hillenbrand et al. (1995, p. 3108) included them in their study of GAE simple vowels. So, the discussion of [+tense] vowels [i, e, o, u, ɑ] is based entirely on Hillenbrand et al. When the features of tenseness and backness are used in tandem, we see that all tense vowels are perceived

accurately: [i] 99.6%, [e] 98.3, [u] 97.2%, and [o] 99.2%. The only tense vowel that is somewhat less accurately perceived is [ɑ] 92.3%, for reasons that were explained in the previous section. Various studies of the acquisition of English vowels by L2 speakers do not show any pronunciation difficulties with these vowels (Koffi, 2012; Munro & Derwing, 2006). The reason for this is that the tense vowels [i, e, u, o, a] are attested in many languages. In fact, Crothers (1978) notes that [i, u, a] are universal, and [e, o] are near universal.

Overlapping Front Lax Vowels

GAE vowels that have the feature [-tense], i.e., lax vowels, tend to overlap in perceptual space. The English vowels in this category are /ɪ, ɛ, æ, ʊ, ʌ, ɔ/. This group of vowels can be further sub-categorized between front vowels and back vowels. Front lax vowels overlap among themselves, and back lax vowels also overlap among themselves. Since the former overlap is different from the latter, the overlapping patterns must be dealt with separately. This section is devoted to overlapping front lax vowels, and the next will focus on overlapping back lax vowels.

The confusion data shows that [ɪ] never overlaps with [æ]. This means the vowels [ɪ] and [æ] are perceptually distant. This makes sense since [ɪ] is a high vowel, and [æ] is a low vowel, we expect them to be perceptually different. According to Peterson and Barney's (1952) confusion data, the vowel that straddles both [ɪ] and [æ] is the mid vowel [ɛ]. It is confused with [æ] 9.2% of the time for Perceived Stimuli, and with [æ] 2.9% of the time for Spoken Stimuli. The total confusion rate between these two segments is 12.1%. The vowel [ɛ] is also confused with [ɪ] 2.5% of the time, and [ɪ] with [ɛ] 6.7% of the time, for a total confusion rate of 9.2%. The perceptual distance between these three vowels is shown in Figure 2:
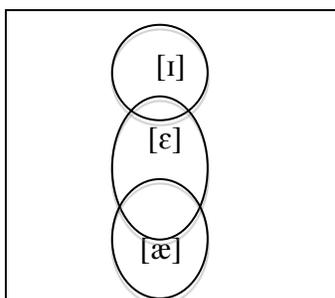
Figure 2. Overlapping between [ɪ, ɛ, æ]

The overlapping between [ɛ] and [ɪ], and [ɛ] and [æ] is phonologically conditioned. The phonological process is one of vowel raising. Some speakers raise [æ] to the height of [ɛ], and [ɛ] to the height of [ɪ] in predictable ways. The exact nature of the conditioning environment varies slightly from region to region, but the raising occurs immediately before voiced consonants. Ladefoged and Disner (2012, pp. 44-45) indicate <bad> [bæd] is pronounced [bɛd] in many of the northern metropolitan areas of the United States (Detroit, Rochester). In Central Minnesota, [æ] is raised to [ɛ] before [g]. In some regions, including the Ohio Valley, parts of Indiana, and in some southern states, [ɛ] is raised to the level of [ɪ] before [n]. So, words like *pen* and *Ben* are homophones with *pin* and *bin*. Lado and Fries (1958, p. 41) write in a note to instructors about teaching the distinction between [ɪ] and [ɛ] to nonnative speakers, saying that, "The distinction between [ɛ] and [ɪ] is likely to be more difficult than the distinction between [e] and [ɛ] and needs special attention."

The relative functional load between [ɪ] and [ɛ] is 54%, and that of [ɛ] and [æ] is 51%. This means that there are many lexical items in English in which these two vowels are contrastive. Consequently, L2 English talkers cannot confuse these pairs of vowels without impinging on intelligibility. GAE hearers expect [ɪ] and [ɛ], and [ɛ] and [æ] not to be confused except in cases where the confusion is predictable by phonological rules. In such environments, distinctions between the phonemes [ɪ] and [ɛ] are neutralized, as are distinctions between [ɛ] and [æ]. Even in such cases, if the hearer is not familiar with the dialect of the talker, intelligibility can occur. At a recent English Department meeting, the faculty representative of the department to the Interfaculty Organization reported that the "[ʃɛl] clauses" in the contract are being renegotiated. Many faculty members were confused and asked what the "[ʃɛl] clauses" were. The misunderstanding was not clarified until the faculty representative spelled the word. It turned out that the contract renegotiation deals with the clauses that have the modal verb "shall" in them. The talker is a native GAE speaker from a different area of the country. In her sociolect, [æ] has risen to [ɛ] before [l]. However, for most people in the audience, this phonological rule does not occur in their sociolect. This unexpected pronunciation of "shall" as [ʃɛl] led to an unintelligible exchange. This confusion confirms that Levis (2011, p. 64) is correct in stating that, "A lot of intelligibility comes down to our expectations. And any time you mess with expectations whether at the phonological level or at the lexical level or at the syntactic level or at the cultural level, you can impair intelligibility." In a more general sense, Rubin (2012, p. 12) echoes the same sentiment by noting that "Notwithstanding society's reliance on speech assessments, it should come as no surprise to discover that such perceptions are highly susceptible to the listener's own expectations of what she is about to hear."

Overlapping Back Lax Vowels

The situation with the back vowels [ʊ, ʌ, ɑ/ɔ] is similar to that of the overlapping front vowels discussed in the previous section. Here, the vowel that overlaps is [ʌ]. According to Peterson and Barney (1952), it is confused with [ɑ/ɔ] 5.2% of the time; [ɑ/ɔ] is confused with it 2.2% of the time. The vowel [ʌ] is confused with [ʊ] 1.6% of the time, and the latter is confused with the former 1.6% of the time. Among Midwesterners, [ʌ] is confused with [ɑ/ɔ] 3.7% of the time, and [ɑ/ɔ] is with [ʌ] 3.3% of the time. As for [ʌ] and [ʊ], the confusion rate in the Midwest is 3.2%, and [ʊ] with [ʌ] is 1%. Indeed, Small (2005, p. 79) notes that, "Students often confuse /ʌ/ with /ʊ/." The acoustic vowel space of Central Minnesota female speakers displayed in Figure 1 shows that [ʌ] is right in the middle of [ʊ] and [ɑ/ɔ]. It explains why [ʌ] is relatively easily confused with both [ʊ] and [ɑ/ɔ]. It has been reported in the L2 phonology literature that [u] and [ʊ] are merging (Celce-Murcia et al., 2010, p. 120). However, this claim is not supported by acoustic data or confusion data. From a standpoint of acoustic phonetics, what is happening is that [ʊ] is undergoing lowering. As a result, it is increasingly being confused with [ʌ]. I have done some informal confusion testing with Dragon Dictation, an automatic speech recognition application. The software does not sufficiently differentiate between <look> vs. <luck>, <book> vs. <buck>, <took> vs. <tuck>, <could> vs. <cud>, <put> vs. <putt>, <stud> vs. <stood> when these words are pronounced by some Central Minnesota speakers. More often than not, the software failed to recognize [ʌ] and perceived it as [ʊ]. I have done the same informal confusion testing with minimal pairs containing [ɑ/ɔ] vs. [ʌ] with the same results. The pairs of words tested were <duck> vs. <dock>, <dug> vs. <dog>, <hut> vs. <hot>, <cut> vs. <cot>/<caught>, <rub> vs. <rob>, <bust> vs. <bossed>, and <hug> vs. <hog> when these words are pronounced by Central Minnesotans. Occasional misunderstandings even among NAE talkers and hearers underscore the confusability of [ʌ]. Recently, the members of my daughter's middle school

swim and dive team were completely confused when they were told that they should pick up their <mums> after practice. They all thought the sentence did not make sense because they did not know how to drive yet. Moreover, it is usually their moms who pick them up after practice, not the other way around.
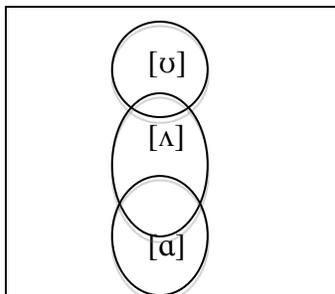

Figure 3. Overlapping between [ʊ, ʌ, ɑ]

Figure 3 provides a visual representation of the overlapping perceptual distance for these three back lax vowels. It underscores the fact that vowels [ɑ/ɔ] and [ʌ] are confused more often than [ʌ] and [ʊ]. Peterson and Barney's data gives a confusion rate of 7.2% for the former, and 3.2% for the latter. Hillenbrand et al. (1995) have a 7% confusion rate for [ɑ] and [ʌ], and a 4.2% rate for [ʊ] and [ʌ]. Catford (1987, p. 90) notes that the relative functional load of [ɑ/ɔ] and [ʌ] is 65%, while that of [ʊ] and [ʌ] is only 9%. Munro and Derwing (2006, p. 493) report that the Mandarin and Slavic participants in their experimental study did not improve their production of [ʊ]. Munro, Flege and MacKay (1996, p. 328) also write that the vowel [ʌ] was the most poorly identified by Canadian hearers who listened to speech samples produced by 240 Italian speakers. These studies indicate that the back vowels [ʊ, ʌ, ɑ/ɔ] are a challenge for L2 English talkers. This challenge is not likely to go away since their NAE (North American English) teachers themselves are increasingly failing to distinguish between them in their speech.

The merger of [æ] and [ɑ] in L2 English

Peterson and Barney found that [æ] was never confused with [ɑ/ɔ], and that [ɑ/ɔ] was confused with [æ] only .18% of the time. Hillebrand et al. also reported that [æ] was confused with [ɑ/ɔ] only .2% of the time, and [ɑ/ɔ] was confused with [æ] .3% of the time. These confusion rates are so small that they are insignificant. Since GAE hearers hardly ever confuse these two low vowels, it suggests that they are perceptually distant. The Perceptual Distance Hypothesis predicts that if L2 English talkers confuse them, intelligibility issues are to be expected. Is this prediction borne out? To answer this question, let's examine my pronunciations of <i-pad> and <i-pod>. I went to an electronic retail store to take a look at <i-pads>[1] when they first came on the market. I told the retail merchant that I was looking for <i-pads> and wanted to take a look at what he had. The retail clerk was a native GAE speaker. He disappeared for a moment and came back loaded with all types of <i-pods>. Why did he misunderstand me? My <i-pad> was misunderstood for <i-pod> because I do not distinguish between [æ] and [ɑ/ɔ] sufficiently well. As a result, there is practically no distance between these two vowels in my acoustic vowel space:

---

[1] The following phonetic/phonological conventions are used throughout the paper: the symbols < > stand for graphemes, / / for phonemes, and [] for phones and allophones.

Figure 4. Koffi's English L2 acoustic vowel space

The F1 values for my [æ], [ʌ], and [ɔ/ɑ][2] are respectively 829 Hz, 793 Hz, and 821 Hz. The 7 Hz difference between my [æ] and [ɑ/ɔ] is below the threshold of human perception. Phoneticians say that humans cannot perceive frequencies that are below 20 Hz. Even the 36 Hz difference between my [æ] and [ʌ] is perceptually insignificant. My F2 values for the same vowels are respectively 1652 Hz, 1419 Hz, and 1606 Hz. As has been noted repeatedly in this paper, for two vowels not to be confused, an F2 frequency difference of 400 Hz is optimal. As for my [æ] and [ɑ], the acoustic vowel space shows clearly that they clustered together. This means that the vowels in my <pad>, <pod>, and <putt> are acoustically indistinguishable from one another. According to Munro et al. (1996, p. 328), I'm not the only L2 English talker who confuses these three vowels. They found that their NAE judges had trouble perceiving [ʌ] accurately when it was produced by Italian speakers, "Only one vowel could be said to be poorly identified overall: in the case of [ʌ], only 25% of the tokens were identified correctly. … The majority of

---

[2] My [ɑ] and [ɔ] are acoustically distinct. However, since many NAE speakers do not distinguish between the two, they mishear my [ɔ] as an [ɑ]. Circles are drawn around my vowels.

misidentified [ʌ] tokens were heard as [æ], though some were also heard as [ɑ]." Munro and Derwing (2006, pp. 491-2) also found that Mandarin and Slavic speakers had difficulty producing [æ]. They say that "although some /æ/ tokens were incorrectly produced as /ɛ/, a large number of incorrect productions from both speaker groups [i.e., Mandarin and Slavic speakers] fell into other categories. The majority of these cases were heard as /ɑ/ and /ʌ/." In other words, other non-native speakers find it challenging to produce [æ, ɑ, ʌ]. In this case, production difficulties lead to perception difficulties because these vowels have high frequency and moderately high functional load. Catford (1987, p. 89) ranks the relative functional load of [æ] vs. [ɑ]/[ɔ] at 76%, that of [æ] vs. [ʌ] at 68%, and [ʌ] vs. [ɑ]/[ɔ] at 65%. As for their frequency, Faircloth and Faircloth (1973, p. 57) list [æ] as the third most frequent vowel in English.

The combination of high frequency and high functional load makes the vowel /æ/ particularly troublesome for L2 English intelligibility. Munro and Derwing (2006, p. 491) write that /æ/ was the vowel that both Mandarin and Slavic speakers had the most trouble producing accurately. The worldwide distribution of [æ] may explain why speakers from various language backgrounds have a hard time producing it accurately. Cutler, Weber, Smits and Cooper (2004, p. 3675) found that many Dutch hearers could not distinguish [æ] from [ɑ] when they listened to American English. The confusion rate was 17.4%. Similarly, Lecumberri and Iragui (1997, p. 59) found that Spaniards confused [æ] with [ɑ] when they listened to British talkers. Data such as these confirm the claim that L2 English speakers cannot produce accurately phones that they cannot perceive accurately. Crothers (1978, p. 95) helps explain why [æ] is such a difficult phoneme to perceive and produce. He studied 209 languages in the Stanford Phonology Archiving Project. He found that only about 1% of the languages in the archive have both a low front [æ] and a low back [ɑ]. For the vast majority of these languages, there is only one low central vowel /a/. Acoustically, this low central /a/ is closer to /ɑ/ than it is to /æ/. Consequently, /æ/ is often produced in ways that it is perceived by NAE hearers either as [ʌ] or [ɔ/ɑ], but rarely as [æ].

This paper would be incomplete if I failed to make a passing remark about the intelligibility of [ɪ] and [ʊ]. These two vowels are often mentioned as being particularly difficult for L2 English speakers to produce. This may well be the case. However, exploratory L2 English acoustic vowel space data that I have collected over the years does not support this commonly held view. In fact, a comparison between Figures 1 and 4 shows that my [ɪ] is almost identical with the [ɪ] in Peterson and Barney (1952) whereas that of Central Minnesotans deviates significantly from NAE norms. Moreover, my [ʊ] is acoustically similar to the [ʊ] produced in Central Minnesota. My preliminary findings agree with what Munro and Derwing (2006, p. 488) found in their longitudinal study of vowel acquisition. They show that their participants produced /ɪ/ reasonably well over time. A comparison with Munro et al. (1996, p. 315) and Munro and Derwing (2006, p. 493) give conflicting results about [ʊ]. In an earlier study, Munro and his co-authors found that, "A larger number of Italian talkers produced [ʊ] tokens with native-like spectral and temporal properties." However, in a later study, he and Derwing found that their Mandarin and Slavic speakers did not show evidence of a noticeable improvement over time for /ʊ/. The acoustic vowel space data that I have of Mandarin and Slavic speakers do not show that they have any difficulty producing /ʊ/. Even if this were not the case, it is unlikely that their production of [ʊ] would result in unintelligibility. If they were to substitute [u] for [ʊ], that would not cause any intelligibility problem because the relative functional load between the two vowels is only 9%. Furthermore, Peterson and Barney (1952) and Hillenbrand et al. (1995) show

that even NAE talkers and hearers often confuse the same two sounds. Consequently, even if nonnative speakers were to substitute [u] for [ʊ], NAE hearers would not be taken off guard.

SUMMARY

Three important observations can be drawn from the analysis presented in this paper. First, the Perceptual Distance Hypothesis helps predict accurately which L2 English sounds are likely to be unintelligible to NAE hearers. Confusion data suggests that six of the 11 phonemic vowels have overlapping perceptual distance. The vowels in question are /ɪ, ɛ, æ, ʊ, ʌ, ɑ/ɔ/. Coincidentally, these vowels are lax. L2 English phonologists have known for a long time that lax vowels are harder for nonnative speakers to acquire than tense vowels. Secondly, formant measurements, segment frequency, and functional load help situate L2 English intelligibility within the wider context of speech recognition. Thirdly, confusion data that has been accumulated over half a century lends empirical support to L2 English intelligibility research. The Perceptual Distance Hypothesis explains why the vowels [æ, ʌ, ɑ/ɔ] are more prone to unintelligibility when uttered by nonnative speakers of English.

ABOUT THE AUTHOR

Ettien Koffi is Professor of Linguistics. He teaches the linguistic courses in the TESOL/Applied Linguistics M.A. program in the English Department at Saint Cloud State University, MN. He has written three linguistic books and is the author of many peer-reviewed articles on various topics in linguistics. His primary area of specialization is at the interface between acoustic phonetics and phonology. He also has extensive experience in emergent orthographies and in the acoustic phonetic and phonological description of dialect variation. He can be reached via email at: enkoffi@stcloudstate.edu.

REFERENCES

Baart, J. (2010). *A field manual of acoustic phonetics.* Dallas, Texas: SIL International.

Catford, J. C. (1987). Phonetics and the teaching of pronunciation. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory*. (pp. 87-100). Washington, D.C.: Teachers of English to Speakers of Other Languages.

Celce-Murcia, M., Brinton, D. & Goodwin, J., with Griner, B. (2010). *Teaching pronunciation: A reference for teachers of English to speakers of other languages, 2ⁿᵈ Ed.*, New York: Cambridge University Press.

Crothers, J. (1978). Typology and universals of vowel systems. In J. Greenberg (Ed.) *The universals of human language: Volume 2, phonology* (pp. 93-152). Stanford, CA: Stanford University Press.

Culter, A., Weber A, Smits, R., & Cooper N. (2004). Patterns of English phoneme confusions by native and non-native speakers. *Journal of the Acoustical Society of America, 116*, 3668-3678.

Dauer, R. M. (2005). The lingua franca core: A new model for pronunciation instruction? *TESOL Quarterly, 39* (3), 543-550.

Faircloth, S. & Faircloth, M. (1973). *Phonetic science: A program of instruction.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Ferrand, C. (2007). *Speech science: An integrated approach to theory and clinical practice*, *2ⁿᵈ Ed.* New York: Pearson Allyn And Bacon.

Hillenbrand, J., Getty, L.A, Clark, Michael J., & Wheeler K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, *97* (5), 3099-3111.

Johnson, K. (2012). *Acoustic and auditory phonetics*, *3rd ed.,* Malden, MA: Wiley-Blackwell.

Jun, H. G., & Li, J. (2010). Factors in raters' perception of comprehnsibility and accentedness. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, Iowa State, University (pp. 53-66). Ames, IA: Iowa State University.

Koffi, E. (2012). Acoustic phonetic self-discovery: The editor's prologue. *Linguistic Portfolios 1*, 1-4.

Ladefoged, P. (2006). *A course in phonetics*, 5th edition. New York: Thomson-Wadsworth.

Ladefoged, P. & Disner S. F. (2012). *Vowels and consonants, 3rd ed.,* Malden, MA: Wiley Blackwell.

Lado, R. & Fries, C.C. (1954). *English Pronunciation: Exercises in Sound Segments, Intonation, and Rhythm*. Ann Arbor, MI: University of Michigan.

Lecumberri, L. G., & Iragui, J. C. (1997). L2 perception of English vowels: Testing the validity of Kuhl's prototypes. *Revista Alicantina de Estudios Ingleses, 10,* 55-68.

Levis, J. M. (2005). Changing context and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(5), 369-377.

Levis, J. M. (2011). Assessing speech intelligibility: Experts listen to two students. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Iowa State, University (pp. 56-69). Ames, IA: Iowa State University.

Miller, G. A., & Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America*, *27* (2), 338-352.

Munro, M., & Derwing T. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*, 520-531.

Munro, M., Flege, J., & MacKay, I. (1996). The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics, 17, 313-334*.

Peterson, G., & Barney, L. (1952). Control methods in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 176-84.

Rubin, D. L. (2012). The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgments and decisions. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, Iowa State, University (pp. 11-17). Ames, IA: Iowa State University.

Small. L. H. (2005). *Fundamentals of phonetics: A practical guide for students, 2nd Edition*. New York: Pearson Education, Inc.

PHONETIC IMITATION OF L2 VOWELS IN A RAPID SHADOWING TASK

Arkadiusz Rojczyk, University of Silesia

The current study investigates the production of L2 vowels in rapid shadowing task. A number of studies have demonstrated that talkers converge with the model on a variety of acoustic properties as a result of imitative tendencies in humans. Such tendencies should be also observed in second-language speech in which acquisition of new sound categories results from efficient imitation of nonnative articulatory patterns. Twenty-two Polish learners of English produced tokens of the English low front vowel /æ/ in word-list reading and immediate imitation of the model. This vowel is reported to be difficult to acquire for Polish learners because it can be accommodated by two Polish neighbouring vowels /e/ and /a/. The magnitude of convergence with the model productions of /æ/ was expressed in Euclidean distance values. The results reveal that participants significantly modified their productions as a result of exposure to the model and that they diverged from their articulatory habits shaped by the influence of L1 vowel categories.

INTRODUCTION

Human beings have an inborn ability to imitate a wide range of actions and intentions (Hauser, 1996; Honorof, Weihing, & Fowler, 2011; Nagell, Olguin, & Tomasello, 1993; Whiten & Custance, 1996). This imitative tendency begins immediately after birth (Meltzoff & Moore, 1999) and continues into adulthood (McHugo, Lanzetta, Sullivan, Masters, & Englis, 1985). Speech appears to be a human activity in which imitation is most likely to play a significant role. Children acquire language from their caretakers and peers (Chambers, 1992; Payne, 1980). Adults acquire elements of the new dialect after moving to a new area (Delvaux & Soquet, 2007; Evans & Iverson, 2007; Munro, Derwing & Flege, 1999; Trudgill, 1986). All this points to the conclusion that language users constantly interact with and imitate patterns occurring in the ambient language.

Sources of such imitative tendencies among speakers are explained from different perspectives relating to human behaviour and cognition. Sociolinguistic theories such as Communication Accommodation Theory (Shepard, Giles, & Le Poire, 2001) assume that individuals accommodate speech features of interacting partners in order to manipulate social distance. Accordingly speakers can both converge with and diverge from interacting partners by subconscious manipulation of attributes such as accent, speaking rate, intensity, utterance duration and frequency of pauses (Giles, Coupland & Coupland, 1991; Gregory & Webster 1996). Meltzoff and Moore (1999) suggest that imitation allows infants to develop the view of self as part of social cognition built on reciprocal imitation of other people. Finally, neurological accounts ascribe imitative tendencies to the architecture of mirror neurons in the human brain (Arbib & Rizzolatti, 1997).

Phonetic imitation (also phonetic convergence or phonetic accommodation) is the process by which a talker takes on acoustic characteristics of the individual that he or she is interacting with (Babel, 2012). This interaction is captured by exemplar-based models (Hintzman, 1986; Nosofsky, 1986), which assume that detailed information in the speech is preserved as exemplars that form a perceptual category. For example, Pierrehumbert (2006) argues that speech production and perception are not, as traditionally viewed, modular but rather that allophonic details as well as speaker information are actively communicated both in production and perception. Such imitative processes are especially important in second-language speech, which is characterised by strong and complex influences from native sound categories on target L2 categories (e.g., Best, 1995; Best & Tyler, 2007; Escudero & Boersma, 2004; Flege, 1987; 1995). Only effective imitation of nonnative properties will lead to formation of new sound categories. The current study investigates how and to what extent imitation in rapid shadowing after the model speech can lead to the production of more native-like vowels. Immediate imitation in shadowing is characterised by a minimum time-lag between hearing the model and actual imitation. This paradigm should be most conducive to attaining approximation of target formant frequencies of L2 vowels because the auditory input is immediately fed to imitative production. In other words, episodic traces of perceived model speech will be reflected in production (Goldinger, 1996, 1998). Moreover, the specificity of the task itself, in which learners are instructed to imitate the model speech without reference to semantics of words, is captured by phonetic as opposed to phonemic perception (Werker & Logan 1985). The phonetic perceptual mode is sensitive to allophonic variation as well as acoustic properties which are absent in the native language.

IMITATION OF VOWELS

Many studies have reported the influence of imitated model speech on production of fine-grained speech properties. Shockley, Sabadini and Fowler (2004) reported that talkers imitate lengthened VOT values for voiceless /p, t, k/ in English. Nielsen (2011) expanded on this observation by showing that longer VOTs as a result of imitation are generalized to new instances of the target phoneme. Most recently, Rojczyk (2012) showed that imitation of VOT is also observed in talkers whose native language does not exploit long VOT values. Honorof et al. (2011) found imitative convergence with the model speech for different degrees of velarization of /l/, measured as the distance between F2 and F1.

A number of studies have found imitation of vowels understood as a reduced acoustic and perceptual distance between baseline to shadowed tokens. Most of these studies conclude that degree of such convergence may depend on both characteristics of the model as well as on which vowels are imitated. Babel (2010, 2012) reported that such convergence of vowels may be selectively modulated by implicit attitudes towards race and nationality of the model. Pardo (2010) and Pardo, Cajori and Krauss (2010) observed that vowel quality is a factor in imitation studies. Talkers may converge, diverge, or not change on some vowels. This tendency was later confirmed in a long-term exposure study on phonetic convergence in college roommates (Pardo, Gibbons, Suppes & Krauss, 2012). Babel (2012), in a lexical shadowing task, observed a greater tendency to imitate low vowels relative to /i/ or /u/. Most importantly for the current study, the vowel /æ/ exhibited the greatest imitative effect. While Babel (2012) ascribed this effect to greater regional variation of low /æ/ and /ɑ/ in American English, another explanation may be formulated by referring to articulatory specification of low and back vowels. Low vowels, unlike high vowels, are characterized by greater mouth opening and jaw lowering, which leaves more

space for individual variability in their production. Such variability will contribute to more pronounced convergence effects observed in imitation.

THE CURRENT STUDY

The current study examines imitation of the English vowel /æ/ by Polish learners. This vowel is commonly reported to be one of the most difficult to acquire by nonnative learners of English (Bohn & Flege, 1997; Flege, Bohn & Jang, 1997; Strange, Akahane-Yamada, Kubo, Trent, & Nishi, 1997) and to be a marker of foreign-accentedness (Flege, 1992; Major, 1987). Polish learners of English, whose native language does not have low front vowels (Jassem, 2003), have difficulties with establishing a new vowel category for /æ/ (Gonet, Szpyra-Kozłowska and Święciński, 2010; Rojczyk, 2011; Sobkowiak, 2003). Applying the assimilatory metric, English /æ/ is equally likely to be assimilated by front mid /e/ and low central /a/ in Polish. However, the direction of assimilation may depend on many factors ranging from personal preferences (Sobkowiak, 2003) to spelling convention (Gonet et al., 2010).

The major goal is thus to investigate if and to what degree imitation in immediate shadowing will allow Polish learners to approximate target-like formant frequencies of nonnative vowel /æ/. As previously reported, this vowels provides the greatest imitative effect in imitation by native speakers (Babel, 2012); however it is not known if and to what extent this vowel will be imitated by talkers with a different language background. In order to quantify the imitative convergence in this scenario, formant frequencies of /æ/ vowels were compared between two tasks: word-list reading (baseline condition) and shadowing after the model voice. The metric of imitation was calculated as the Euclidean distance of individual productions in the two tasks to the model productions to reveal a change as a result of auditory exposure to the model talker (Babel, 2012). Lower Euclidean distance values in the shadowing task are expected to show the degree of convergence with the model and, accordingly, the articulatory approximation towards a nonnative vowel category. Moreover, gender will be incorporated in the statistical model as an independent variable, because of previous reports suggesting that gender may be a factor in the magnitude of imitation (Pardo, 2006).

Participants

Twenty-two native speakers of Polish (sixteen females and six males) were included in the study. All of them were recruited from the University of Silesia in Poland. Their mean age was 19.8 (*SD* = .03). Their self-reported proficiency in English ranged from intermediate to upper-intermediate. None of the participants reported any speech or hearing disorders.

Materials

The words used in the experiment were twelve monosyllabic sequences with the vowel /æ/ flanked by consonants (Table 1). They were recorded for the shadowing task by a male southern British English speaker using the recording equipment reported below. The model talker was instructed to use natural speaking tempo and falling intonation for each token. Each model vowel was measured as described below to obtain F1 and F2 formant frequencies of /æ/s in each token used for shadowing. The raw model values for /æ/ in each word are provided in the Table 1.

Table 1

Words Used in the Experiment with the Model Talker's Frequencies of the First and Second Formant Expressed in Hz

Word	F1	F2
back	749	1492
bad	697	1558
bat	683	1570
cab	696	1618
cap	785	1631
cat	688	1620
dad	706	1675
fat	802	1544
hat	676	1593
sad	720	1641
pack	673	1575
mad	727	1594

Procedure and Recording

The experiment took place in the Acoustic Laboratory at the Institute of English, University of Silesia. Data were collected in two blocks. The first block was reading the list of words to establish baseline productions of /æ/. The participants were instructed to read the words using natural intonation and articulatory rate. The words were presented sequentially on a monitor screen in 54-point black font in the middle of the screen. Twelve other foil words with different vowels were randomly dispersed among target words to distract the talkers' attention from the object of the experiment. The second block was used for immediate shadowing after the model talker. The participants were instructed that upon hearing a word spoken by the voice they were to immediately repeat it. The presentation of words was separated by a two-second interval after the cessation of imitations. Five foils were used at the beginning of this block to familiarize the participants with the procedure. At the end of the session, the participants read /bVt/ sequences with Polish vowels /i, e, a, o, u/ that were further used as landmark points to establish the acoustic space for each talker in normalization. Each session lasted approximately twenty minutes.

The recordings were made in a sound-proof booth, the signal was captured with a headset dynamic microphone Sennheiser HMD 26, preamplified with USBPre2 (Sound Devices), into .wav format with the sampling rate 48 kHz, 24 bit quantization. The model voice was provided by high quality headphones built in the headset.

Measurements

Formant frequencies of vowels were measured at vowel midpoint using the add-on vowel analysis software Akustyk 1.8 (Plichta, 2011) for Praat (Boersma, 2001). First, all recordings were downsampled to 10 kHz and vowel midpoint was located using wideband spectrograms. Formants were tracked using a 25-ms Hanning window with default 11 (female) and 12 (male) poles. If the tracker yielded spurious or missed formants, LPC spectral envelopes and FFT power spectra were compared in order to recompute a prediction order so that it would match a particular speaker's voice quality. The total number of measured target tokens was 528 (22 talkers x 24 vowels). In order to compare the distance of individual productions to model production, anatomical and physiological variation between talkers was normalized using the Lobanov transform (Lobanov, 1971; Adank, Smits & van Hout, 2004).

Analysis

In order to calculate how much participants modified their production as a result of exposure to the model production, the Euclidean distance was computed between the participants' and model's F1 and F2 frequencies. This shows how far apart individual vowels are on a vowel plane. As a result, the magnitude of the convergence was expressed in the distance values. In this metric, the lower the value the more similar the model and participants' values are in the acoustic space. The calculated distances in the word-list and shadowing conditions were used as repeated-measures dependent variables. Data were analysed using a two-way mixed ANOVA with task (word-list, shadowing) and gender (male, female) as independent variables. The dependent variable was Euclidean distance expressed in Hertz. Data for each word out of twelve stimulus words were used as separate rows in the analysis, which yielded 264 rows (22 participants x 12 words). Moreover, scatter plots for individual productions were used to inspect the clustering of participants' vowels with the model vowels.

RESULTS

Figure 1 shows the scattering of individual productions of /æ/ in word-list (black) and imitation (green) around the model production (red). It is evident that shadowed productions are more centered around the model. Unlike vowels from the word-list reading, they are also characterized by less extreme productions towards either Polish /e/ or /a/. This demonstrates that even participants who completely accommodated English /æ/ to either /e/ or /a/ in their native language reacted to the auditory input and modified their productions towards the model vowel. Moreover, the model auditory input generated a magnet effect by cancelling less extremely outlying productions in the imitation task, as demonstrated by better clustering of individual productions around the model in shadowing.
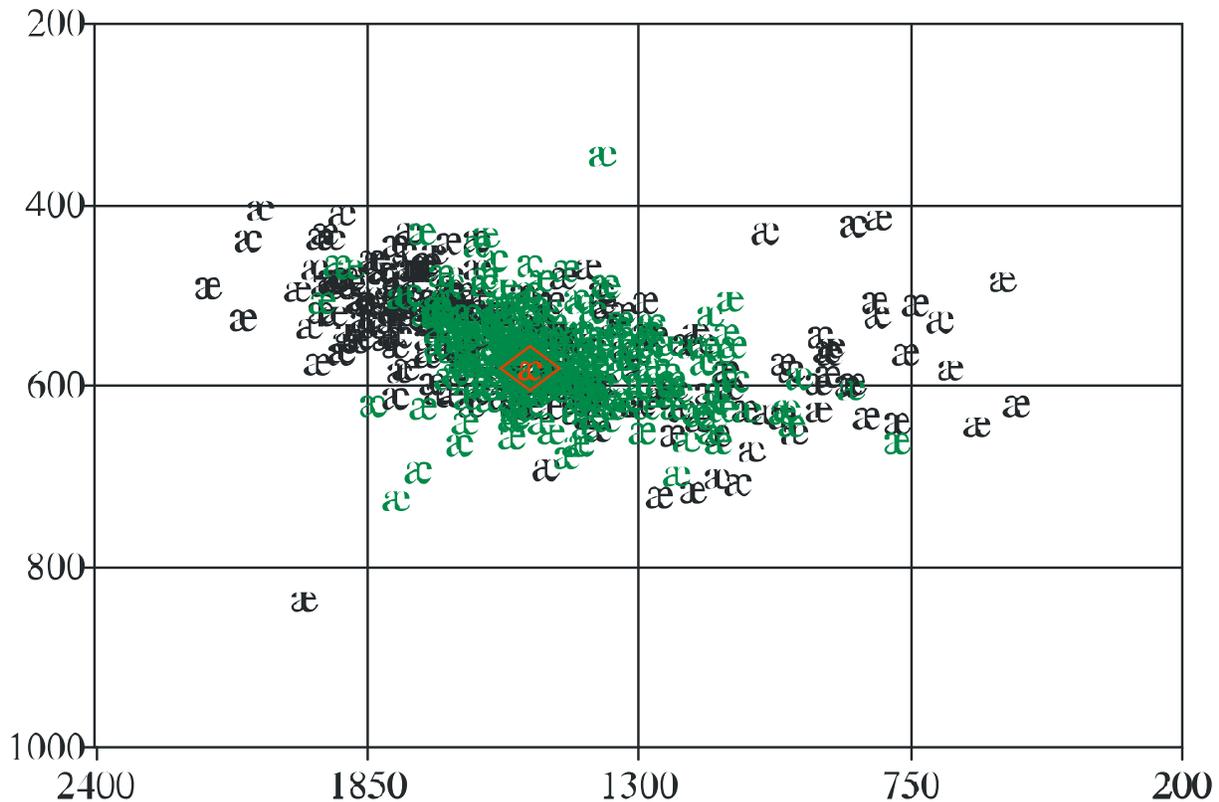
Figure 1. Scatter plot of vowels from read words (black) and imitated words (green) on an F1-F2 plane. Model vowels in a red diamond.

The analysis of Euclidean distances of individual productions to the model vowels in the two tasks revealed a highly significant main effect of task on the magnitude of convergence [F(1, 262) = 43.35, p < .001]. The participants modified the productions of the /æ/ vowels to approximate the model in imitation (*M* = 165; *SD* = 120) compared to baseline word reading (*M* = 264; *SD* = 199). The was no significant gender x task interaction [F(1, 262) = .11, p > .05], indicating that gender of the participants did not affect the magnitude of convergence.

DISCUSSION

The study investigated to what extent nonnative vowels can be imitated in a shadowing task. The degree of imitation was calculated as the Euclidean distance of individual productions to the model vowels. In order to assess the magnitude of imitation, the productions from shadowing were compared to baseline reading of words for each participant. The vowel was low front /æ/ in English, which is difficult to acquire for Polish learners who accommodate it in production and perception to neighbouring /e/ and /a/.

The results revealed a significant convergence with the model in the task in which talkers were required to immediately repeat after the model voice compared to the task in which they read orthographic representations of the words. Accordingly, it suggests that foreign language learners are able to modify their productions of nonnative vowels as a result of exposure to the model. This is confirmed by significantly lower Euclidean distance values in the shadowing task. If /æ/ tokens from the word list are taken to represent participants' default exemplars of this

vowel, the tokens from imitation show that learners' vowel categories are not without exception shaped by L1 categories. Obviously, the time-course of such convergence is probably limited, in that in order for a learner to modify their vowel production, the interval between exposure and the onset of imitation must be relatively short. This is suggested by research with nonnative imitation in immediate and distracted tasks (Rojczyk, 2012). In this study, Polish learners produced tokens with voiceless plosives in English and their VOT was measured. Polish, unlike English, does not use long-lag VOT for /p, t, k/ and, as a result, Polish learners have difficulties producing sufficiently long VOT values in English. Participants' VOT was measured in voiceless plosives in the word list, immediate and distracted imitation. In the distracted task learners were required to listen to the model, read the number on the screen, and then begin imitation. The results revealed that VOT values in this task were intermediate between baseline word-list reading and imitation, indicating that if the interval between exposure and imitation is lengthened or cognitively taxing, learners resort to their habitual production patterns. The same regularity may be expected to occur for vowel production, in that if participants are distracted or delayed in their imitation, they will produce tokens which diverge from the model vowels.

The current study did not find any influence of gender on the magnitude of convergence. Such possibility was suggested in previous studies (Pardo, 2006). There are two reasons why this may be the case. First, in the current study male participants were significantly underrepresented, which may have affected the results. Second, the study by Pardo (2006) observed gender differences in conversational interaction. Such interactions are characterized by more psychological and sociolinguistic influences which may trigger gender differences. The current study relied to a greater extent on psychoacoustic reactions to the auditory input, which does not necessarily have to be gender specific.

The current results also confirm previous observations that fine-grained phonetic details are not filtered out in speech perception, as demonstrated by plasticity in speech production (e.g., Nielsen 2011; Norris, McQueen & Cutler, 2003; Sancier & Fowler, 1997). If phonetic detail were discarded in production, participants in the current study would not have modified their production as a result of exposure to the model. By extension, this also suggests that L2 learners are able to restrict the assimilatory impact of native sound categories on target L2 categories, at least if the time interval between the model input and the onset of production is relatively short and undistracted. It is thus possible that the interference of native phonological and articulatory patterns is gradient and its magnitude may depend on circumstances and activity that a learner is engaged in.

Finally, yet another aspect that may have contributed to the current findings is the role of orthography. The effects of orthography in second-language pronunciation are evident (Young-Scholten & Archibald, 2000; Silveira, 2007), especially in the case of learners in the current study whose contact with L2 had relied greatly on written material. Since the target vowel /æ/ was orthographically represented by a letter 'a' in all tested words, it may be assumed that this letter may have caused more divergent productions towards vowel /a/ in the reading task. However, this prediction is not fully warranted. The inspection of individual productions in the reading task clearly shows that the distribution of /æ/ was relatively balanced between Polish /e/ and /a/. This reflects previous observations that the production of this vowel by Polish learners is not tied to its representation by a letter 'a' (Gonet et al., 2010). It is for future studies though to demonstrate the extent of influence of orthographic representation on imitation for this vowel by precise manipulations of tested words.

ABOUT THE AUTHOR

Arkadiusz Rojczyk is an Assistant Professor at the University of Silesia in Poland. His research concentrates on production and perception of second language speech, speech analysis and resynthesis. He is currently working on vowel perception and production in second-language speech

arkadiusz.rojczyk@us.edu.pl
Institute of English, University of Silesia
Grota-Roweckiego 5
41-205 Sosnowiec, Poland

REFERENCES

Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America, 116*, 3099-3107.

Arbib, M., & Rizzolatti, G. (1997). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition, 29*, 393-424.

Babel, (2010). Dialect convergence and divergence in New Zealand English. *Language in Society, 39*, 437-456.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics, 40*, 177-189.

Best, C. (1995). A direct realist view of cross-language speech perception: In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 171-204). Baltimore: York Press.

Best, C., & Tyler, M. (2007). Nonnative and second language speech perception: Commonalities and complementarities. In O.-S Bohn & M. Munro (Eds.), *Language experience in second language speech learning. In honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 10*, 341-345.

Bohn, O.-S., & Flege, J. E. (1997). Perception and production of a new vowel category by adult second language learners. In A. James & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 53-73). Berlin: Mouton de Gruyter.

Chambers, J. (1992). Dialect acquisition. *Language, 68*, 673-705.

Delvaux, V., & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica, 64*, 145-173.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26*, 551-585.

Evans, B. G., & Iverson, P. (2007). Plasticity in vowel perception and production: A study of accent change in young adults. *Journal of the Acoustical Society of America* 121, 3814-3826.

Flege, J. E. (1987). The production of new and similar phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15, 47-65.

Flege, J. E. (1992). The intelligibility of English vowels spoken by British and Dutch talkers. In R. Kent (Ed.), *Intelligibility in speech disorders: Theory, measurement, and management* (pp. 157-232). Amsterdam: John Benjamins.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium: York Press.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25*, 437-470.

Giles, H., Coupland, J., & Coupland, N. (1991). *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge: Cambridge University Press.

Goldinger, S. (1996). Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1166-1183.

Goldinger, S. (1998). Echoes or echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251-279.

Gonet, W., Szpyra-Kozłowska, J., & Święciński, R. (2010). Clashes with ashes. In E.Waniek-Klimczak (Ed.), *Issues in accents of English 2: Variability and norm* (pp. 213-232). Newcastle upon Tyne: Cambridge Scholars Publishing.

Gregory, S. W., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status predictions. *Journal of Personality and Social Psychology, 70*, 1231-1240.

Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93*, 411-428.

Honorof, D. N., Weihing, J., & Fowler, C. A. (2011). Articulatory events are imitated under rapid shadowing. *Journal of Phonetics, 39*, 18-38.

Jassem, W. (2003). Illustrations of the IPA: Polish. *Journal of the International Phonetic Association, 33*, 103-107.

Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America, 49*, 606-608.

Major, R. (1987). Phonological similarity, markedness, and rate of L2 acquisition. *Studies in Second Language Acquisition, 9*, 63-82.

McHugo, G., Lanzetta, J., Sullivan, D., Masters, R., & Englis, B. (1985). Emotional reactions to a political leader's expressive displays. *Journal of Personality and Social Psychology, ,* 1513-1529.

Meltzoff, A., & Moore, M. (1999). Persons and representation: Why infant imitation is important for theories of human development. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (pp. 9-35). Cambridge: Cambridge University Press.

Munro, M. J., Derwing, T. M., & Flege, J. E. (1999). Canadians in Alabama: A perceptual study of dialect acquisition in adults. *Journal of Phonetics, 27*, 385-403.

Nagell, K., Olguin, K., & Tomasello, M. (1993). Processes of social learning in tool use of chimpanzees (Pan troglodytes) and human children (Homo sapiens).*Journal of Comparative Psychology, 107*, 174-186.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics, 39*, 132-142.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 7*, 204-238.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America, 119*, 2382-2393.

Pardo, J. S. (2010). Expressing oneself in conversational interacton. In E. Morsella (Ed.), *Expressing oneself/expressing one's self: Communication, cognition, language, and identity* (pp. 183-196). New York: Psychology Pres.

Pardo, J. S., Cajori, J. I., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, and Psychophysics, 72*, 2254-2264.

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics, 40*, 190-197.

Payne, A. C. (1980). Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In W. Labov (Ed.), *Locating language in time and space* (pp. 179-218). New York: Academic Press.

Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics, 34*, 516-530.

Plichta, B. (2011). Akustyk for Praat (Version 1.8) [Computer program]. Retrieved August 16 2011 from http://bartus.org/akustyk/.

Rojczyk, A. (2011). Overreliance on duration in nonnative vowel production and perception: The within lax vowel category contrast. In M. Wrembel, M. Kul & K. Dziubalska-Kołaczyk (Eds.), *Achievements and perspectives in SLA of speech: New Sounds 2010* (pp. 239-250). Frankfurt am Main: Peter Lang.

Rojczyk, A. (2012). Phonetic and phonological mode in second language speech: VOT imitation. Papaer presented at EuroSLA22 - 22nd Annual Conference of the European Second Language Association, Poznań Poland, 5-8 September.

Sancier, M. L. & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*, 421-436.

Shepard, C. A., Giles, H., & Le Poire, B. A. (2001). Communication accommodation theory. In W. P. Robinson & H. Giles (Eds.), *The new handbook of language and social psychology* (pp. 33-56). Chichester: John Wiley & Sons Ltd.

Shockley, K., Sabadini, L., & Fowler, C. A. (2004). Imitation in shadowing words. *Perception and Psychophysics, 66*, 422-429.

Silveira, R. (2007). Investigating the role of orthography in the acquisition of L2 pronunciation: A case study. *New Sounds 2007: Proceedings of the Fifth International Symposium of the Acquisition of Second Language Speech.*

Sobkowiak, W. (2003). *English phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., & Nishi, K. (2001). Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America, 109*, 1691-1704.

Trudgill, P. (1986). *Dialects in contact*. New York: Blackwell Publishing.

Werker, J. F., & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics, 37*, 35-44.

Whiten, A., & Custance, D. M. (1996). Studies of imitation in chimpanzees and children. In C. M. Heyes & B. G. Galef (Eds.), *Social learning in animals: The roots of culture* (pp. 291-318). San Diego: Academic Press.

Young-Scholten, M., & Archibald, J. (2000). Second language syllable structure. In J. Archibald (Ed.), *Second Language Acquisition and Linguistic Theory* (pp. 64-97). Oxford: Blackwell.

Chun, D. M., Jiang, Y., & Ávila, N. (2013). Visualization of tone for learning Mandarin Chinese. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference.* Aug. 2012. (pp. 77-89). Ames, IA: Iowa State University.

VISUALIZATION OF TONE FOR LEARNING MANDARIN CHINESE

Dorothy M. Chun, University of California, Santa Barbara
Yan Jiang, University of California, Santa Barbara
Natalia Ávila, University of California, Santa Barbara

For a tonal language such as Mandarin Chinese, accurate pronunciation of tones is critical to meaning, and research suggests that computer-based programs that allow for visualization of pitch contours are helpful for improving learners' pronunciation (Chun, 1998; Hardison, 2004; Levis & Pickering, 2004). This paper reports on a pilot study using speech analysis software (*Praat*), which allowed L2 Chinese learners first to hear a native speaker of Mandarin say words and phrases while seeing a visual display of the native speaker's pitch curves, then to record themselves reading the same words and phrases, and later to compare their own pitch contours to those of the native speaker. Students in first-year Chinese were recorded reading words and phrases before and after two computer-based training sessions. Native speakers rated the words and phrases for accuracy of tones. Results indicate a ceiling effect for the pronunciation of mono- and disyllabic word tones, with 83.39% of tones produced correctly in the pre-test. Of the 16.41% of tones that were incorrect in the pre-test, almost 50% of them were pronounced correctly in the post-test. Students indicated in a post-study survey that seeing the pitch curves of both the native speakers and their own helped them improve their word tones.

L2 PRODUCTION OF MANDARIN CHINESE TONES

In Mandarin Chinese, there are 4 lexical tones manifested by different *pitch levels* (in acoustic phonetic terms, *fundamental frequency* or *F0*) and different *pitch contours*. As can be seen in Figure 1, which contains tones spoken by a native Mandarin speaker, tone 1 is high and level; tone 2 has rising pitch; tone 3 falls and then rises; and tone 4 starts high and falls sharply. A neutral tone that is unstressed and much weaker in intensity and shorter in duration also exists.
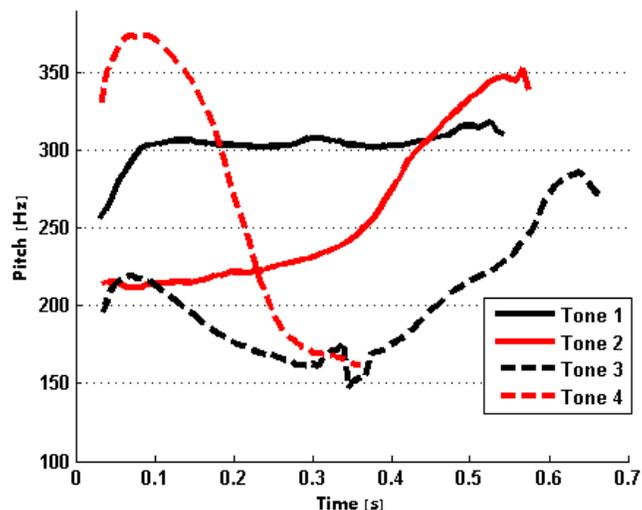

Figure 1. Acoustic Representation of Four Mandarin Chinese Tones

Research on the L2 acquisition of Mandarin Chinese tones has shown that learners whose L1 is a non-tonal language, e.g., English, have difficulties in perceiving and producing tones, but that training in both perception and production can be effective (Wang, Sereno, & Jongman, 2006). There is not universal agreement with regard to which tones are the most difficult. Shen (1989), for example, found that tonal errors made by American learners who had studied Chinese for four months ranged from 8.9% for tone 2 to 55.6% for tone 4, and that the mispronunciation of tones tends to lie in the pitch height (low, middle, or high) and not in the pitch contour (level, rising, dipping, or falling). For second-year American learners, Miracle (1989) found an overall error rate of 42.9%, with the errors relatively evenly divided among the tones. Chen (1997) found that tones 1 and 4 were easier than tones 2 and 3.

Molholt and Hwu (2008) advocated using acoustic analyses of speech (e.g., Figure 1) rather than simple descriptions or graphic representations that have traditionally been used in instructional materials (Figure 2). In line with this approach, other studies of L2 learning of non-tonal languages have demonstrated that when learners see acoustic visualizations of native speakers compared with their own pitch curves, the learners improve their pronunciation (Hardison, 2004).
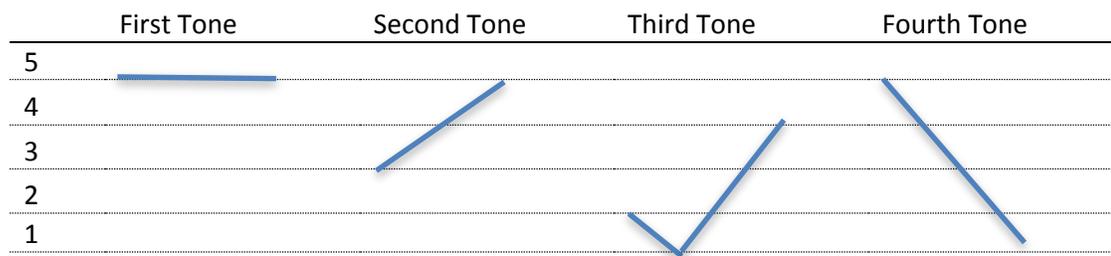
Figure 2. Typical Depiction of 4 Mandarin Tones.

PRESENT STUDY

This pilot study sought to determine whether providing L2 learners with visualizations of Mandarin native speakers' tonal pitch contours along with the learners' own pitch contours would improve L2 production of tones.

Participants

Sixteen students who were in their first-year (third-quarter) of studying Mandarin Chinese at a large state university in the western U.S. volunteered to participate. A list of the questions in the demographic survey can be found in Appendix A. None of them was a heritage speaker of Mandarin.

Materials and Procedures

The study was conducted in two one-hour sessions, a week apart. In the first session, students were given a pre-test: they were asked to read a list of 31 monosyllabic and disyllabic words or phrases and were recorded. They were then seated at a computer and shown the pitch curves of a Mandarin native speaker saying the 31 words or phrases (see Figure 3) and simultaneously heard the words spoken as they viewed the pitch curves that had been created with the acoustic software *Praat* (freely available at http://www.fon.hum.uva.nl/praat/). For 30 minutes, students

practiced repeating the words and phrases while hearing the native speaker and viewing the pitch curves.
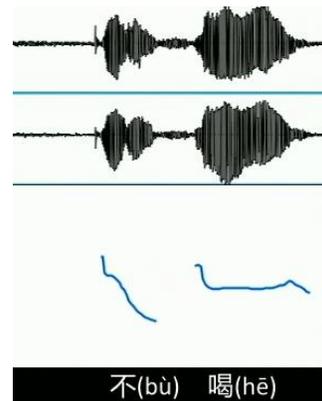
Figure 3. Native Speaker's Pitch Curves *bù hē* ('not drink').

A week later, students returned to the computer lab. The words that they had read the previous week had been pitch tracked by the researchers, and students could once again view the native speaker's pitch curves, but they could now also view their own production of these words and compare the two pitch curves (Figure 4).[1] After practicing again for 30 minutes, they were recorded reading the original 31 words and phrases. This recording would serve as the post-test; Figure 5 shows a student's pre- and post-test pronunciation of *bù hē*.
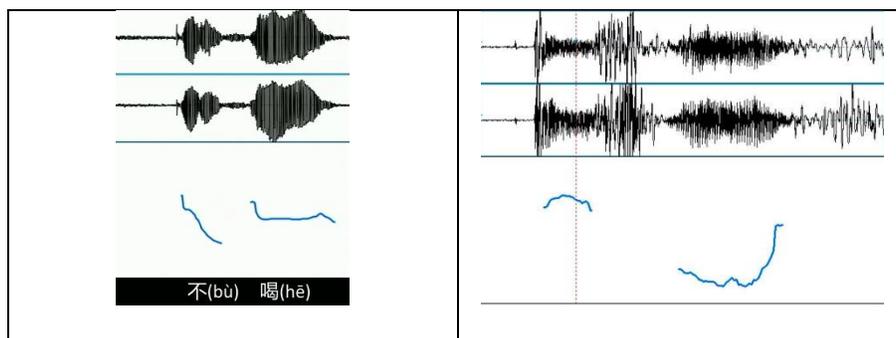
Figure 4. Native Speaker's *bù hē* (left) and Student's Pre-test *bù hē* (right).

---

[1] In other studies that used visualization of learners' pitch curves (e.g., Hardison, 2004), commercial software that provided a display of pitch curves immediately after learners' production was used. However, our study used the open source software *Praat*, which does not automatically provide pitch curves but rather requires student recordings to be processed individually "by hand." It is to be sure not as ideal as automatic, instantaneous feedback, but it does not require the purchase of costly software.
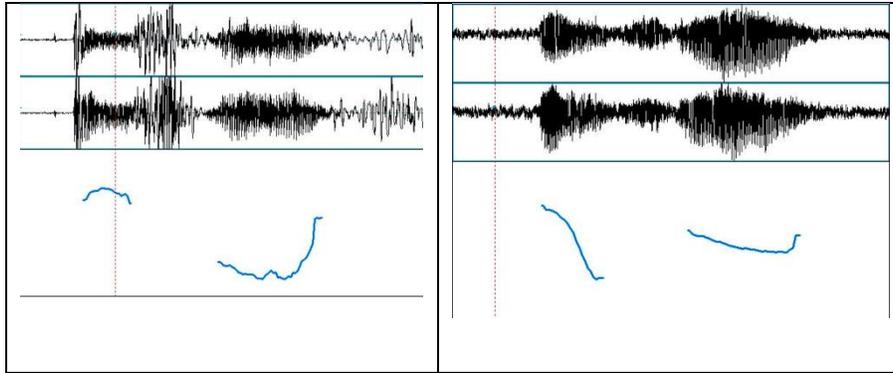
Figure 5. Student's Pre-test *bù hē* (left) and Student's Post-test *bù hē* (right).

Data Analysis

The pre- and post-tests were then assessed in two different ways in order to determine whether improvement had occurred. First, four native speakers of Mandarin Chinese (university Teaching Assistants) were presented with audio files of the 31 words spoken by the 16 learners. The words from both pre- and post-test were mixed together and randomized so that the raters would not know whether the recording was from the pre-test or the post-test. Raters assigned scores on a scale of 1-4, with 1=mostly correct and 4=mostly incorrect. Second, the same four native speakers were presented with pairs of words (the same words from the pre-test and post-test respectively) and asked to rate the improvement from pre-test to post-test on a scale of 1-5, with 1=great improvement; 2=some improvement; 3=only a little improvement; 4=no improvement or worse; 5=no change, both correct.

After the second session, students filled out a brief survey about the usefulness of seeing pitch curves, both of the native speakers and of their own.

RESULTS

Nineteen syllables in isolation or combined with others were selected for tone analysis based on a representative sampling of the four tones and tone combinations (Table 1). Among them were four 1st tone, five 2nd tone, six 3rd tone, and four 4th tone syllables. Within each tone group, at least one syllable belongs to each of the four categories: syllable in isolation, disyllabic tone + tone combination (t+t), disyllabic tone + neutral tone combination (t+n), and syllable in a sentence. These tone combinations are very common in Mandarin Chinese, and mastery of them is critical for learners to acquire correct pronunciation. Examining the students' accuracy of producing those syllables allowed us to determine which tones and tone combinations were most problematic for learners. It was also a systematic way to test the effectiveness of using pitch curves for tone training.

Table 1
Selected Syllables (Bolded) for Analysis

	Syllable in isolation	Disyllabic t+t	Disyllabic t+n	Syllable in a sentence
1st tone	***Shū*** 'book'	*hǎo **shū*** 'good book'	***māma*** 'mother'	*zhè shì wǒ bàba **māma*** 'this is my father, mother'
2nd tone	***chá*** 'tea'	***nín** hǎo; hǎo **chá*** 'how are you'; 'good tea	***péng**you* 'friend'	*zhè shì wǒ **péng**you* 'this is my friend'
3rd tone	***nǐ; hǎo*** 'you'; 'good'	***nǐ** hǎo* 'how are you'	***nǐ** men* 'you' (pl.)	***nǐ** hǎo; **nǐ**men hǎo* 'how are you'; 'how are you' (pl.)
4th tone	***kàn*** 'look'	*hǎo **kàn*** 'good looking'	***bà**ba* 'father'	*zhè shì wǒ **bà**ba māma* 'this is my father, mother'

Pre-Test Results

The overall score statistics from the pre-test are displayed in Table 2, with the percentage of each score being the ratio of the score count over the total 304 syllables collected from the 16 participants. The learners' pronunciation of tones was rated "1" (mostly correct) 79.61% to 87.17% of the time among the four raters. The percentage of tones that was rated "4" (mostly incorrect) ranged from 0.99% to 9.87% among the four raters. On average, 83.39% of the tones were rated "1", and only 5.67% were rated "4." In other words, raters perceived that the students mostly produced correct tones, but rarely incorrect tones. This suggests that a "ceiling effect" may exist, i.e., the training may not be effective for learners at this proficiency level, which may be attributed to the fact that they had been learning Chinese for at least six months before taking the test.

Table 2
Percentage (Pct.) of Scores in Pre-Test and Post-Test (1=mostly correct; 4=mostly incorrect)

Score	Avg. Pct. (%)	
	Pre-test	Post-test
1	83.39	85.23
2	7.24	6.93
3	3.70	3.14
4	5.67	4.70

We further broke down the ratings of tones to determine whether the students' proficiency levels on the four tones were different. As shown in Figure 6, the 1st tone had the highest average percentage of rating "1" (90.23%) and the lowest average percentage of rating "4" (2.34%) among the four tones, indicating that the learners had the highest level of proficiency with 1st tone before training. On the other hand, the 4th tone had the highest average percentages of rating "4" (9.38%), suggesting the students made more mistakes when pronouncing 4th tone.
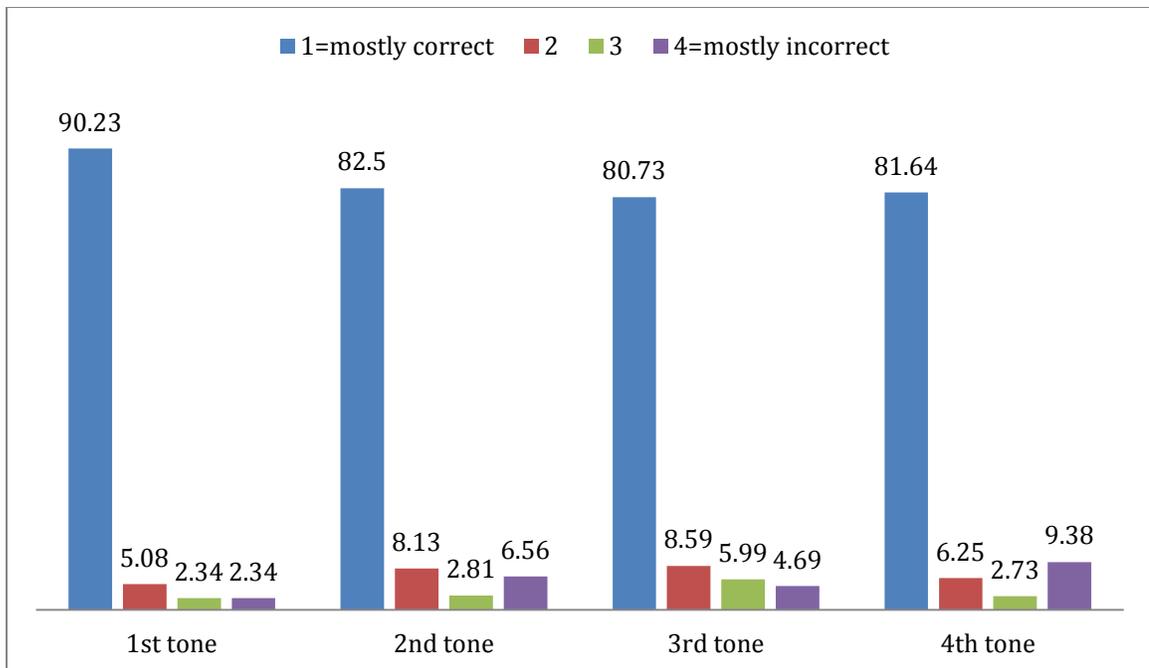
Figure 6. Average Percentages (%) of Each Score on Four Tones in Pre-test.

Post-Test Results

The first method used to determine whether students' tone production improved was to compare the percentage of scores in the post-test to the pre-test. Table 2 shows that the average percentage of score 1 increased by 1.84%, and average percentage of score 4 decreased by 0.97% from pre- to post-test. The change of scores suggests that the students produced more correct tones and fewer incorrect tones after training, although the difference in pre- and post-tests was limited due to the ceiling effect.

To further understand which tones showed improved production after training, we broke down the results by tones and computed the difference of the average percentages of correct tones in post- and pre-tests (see Figure 7). The higher average percentages of correct tones in the post-test indicate that learners improved on 2nd, 3rd and 4th tones after receiving the training. Notwithstanding the ceiling effect, learners improved their 4th tones by 5.42%. The reason might be that learners had a lower proficiency level on 4th tone before training, as suggested by the pre-test results.
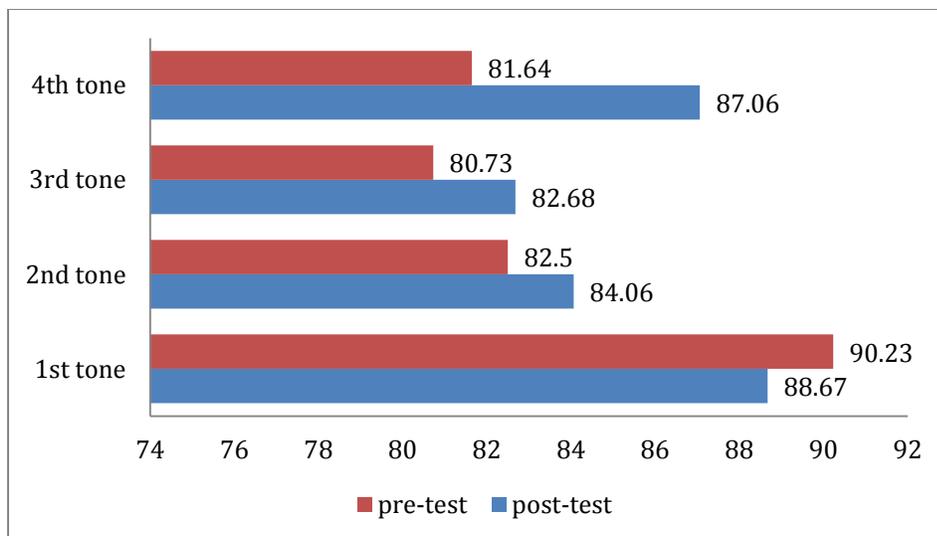
Figure 7. Percentages (%) of Correct Tones in Pre- and Post-tests.

Re-analysis of Pre- and Post-Tests

Acknowledging the unexpectedly high initial scores, we re-analyzed the pre- and post-test results and focused on the 16.61 % of the tones that were not produced correctly (i.e., did not receive a score of "1") in the pre-test (refer to Table 2, e.g., 1=mostly correct 83.39%; 4=mostly incorrect 5.67%). When comparing only the pre- and post-tests for the problematic tones (i.e., those rated "2," "3" or "4"), almost half of them, 49.5%, were rated "1" (mostly correct) in the post-test. The other 50.5% received ratings of "2," "3" or "4" in the post-test (see Figure 8).
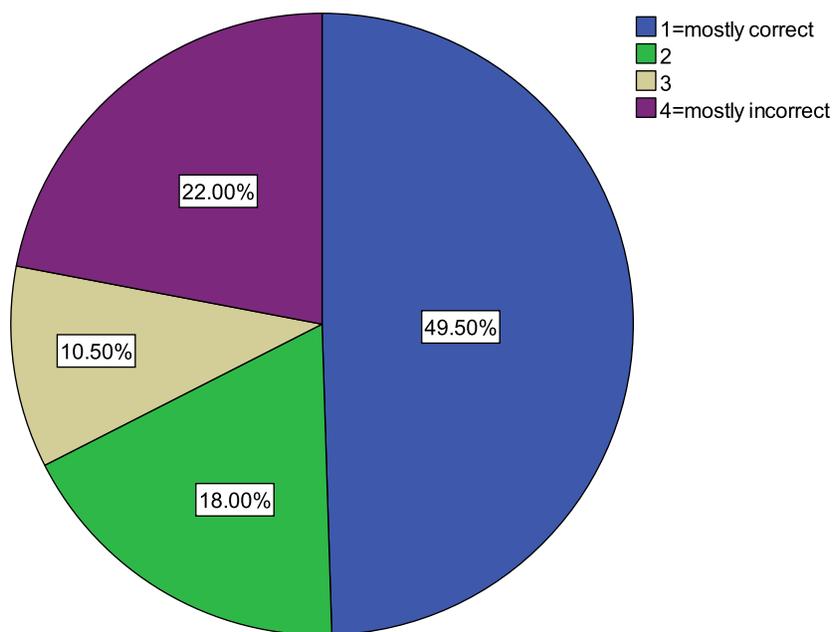
Figure 8. Distribution of Post-test Scores of the Incorrect Tones.

We also performed paired samples t-tests comparing the scores on the pre- vs. post-tests. In comparing all of the data, the mean of the pre-tests ($M = 1.31$, $SD = 0.79$) was not significantly different from the mean of the post-tests ($M = 1.27$, $SD = 0.74$), $t(1211) = 0.44$, $p > .05$. However, the mean scores differed significantly when comparing the ratings of the 16.61% of the tones that were incorrect in the pre-test with the ratings of these tones in the post-test, $t(199) = 10.25$, $p < .001$. The mean score of incorrect tones in the post-tests ($M = 2.05$, $SD = 1.22$) was .85 point lower than the mean score in the pre-tests ($M = 2.90$, $SD = 0.88$). The results showed that students made significant improvement in the incorrect tones with the aid of visualizing pitch curves.

Improvement Results

The second method of determining whether students' tone production improved from the pre- to post-test was to present the native speaker raters with the pre- and post-test recordings, one right after the other. Four native speakers then rated the pre- and post-test pairs from 1=great improvement to 4=no improvement or worse, and 5=no change, both correct. Table 3 shows the improvement results of the 16 students' scores. Due to the strong ceiling effect in the pre-test, the students produced both tones correctly, indicated by score 5, 67.39% of the time. The average of scores 1 to 3 is 17.01%, which means that students made a little, some or great improvement in the post-test approximately 17.01% of the time. Only 15.59% of the pre- and post-test pairs were rated "4," showing "no improvement or worse."

Table 3
Percentage (Pct.) of Scores in Improvement Analysis

Score	1	2	3	4	5
Avg. pct. (%)	5.78	4.36	6.87	15.59	67.39
Avg. of scores 1 to 3			17.01		

We further broke down the results by tones, and computed the means of percentages of scores 1, 2 and 3 (see Figure 9). The average percentages of improvement on 1st, 2nd, 3rd and 4th tone are, 20.24%, 14.29%, 17.91% and 15.87%. The results confirm that students achieved a certain degree of improvement on all four tones about 17% of the time. Tone 2 showed the least improvement and tone 1 the most improvement. Figure 10 is an example of a learner's pitch curves of the 4th tone syllable *kàn* ("to look") in pre- and post-tests. The learner appears to have made significant improvement with the aid of seeing pitch curves. However, Figure 11 is an example of another learner's pitch curves for the same word, but in this case, the pre- and post-tests are similar, both lacking the falling tone contour that is necessary for 4th tone syllables.
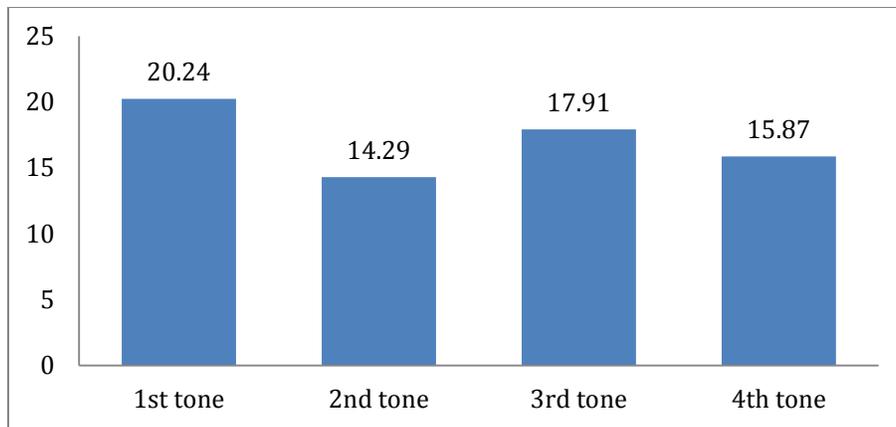
Figure 9. Average Percentage (%) of Improvement for Each Tone.
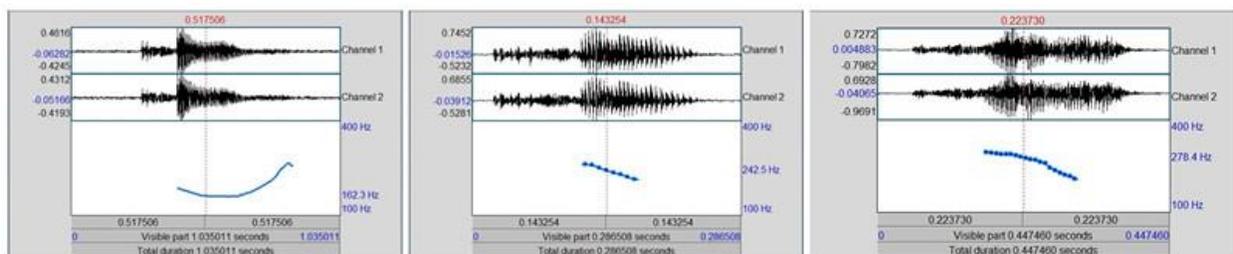
Figure 10. Pitch Curves of *kàn* ("to look") by Learner X in Pre-test (left), Post-test (middle) and by Native Speaker (right).
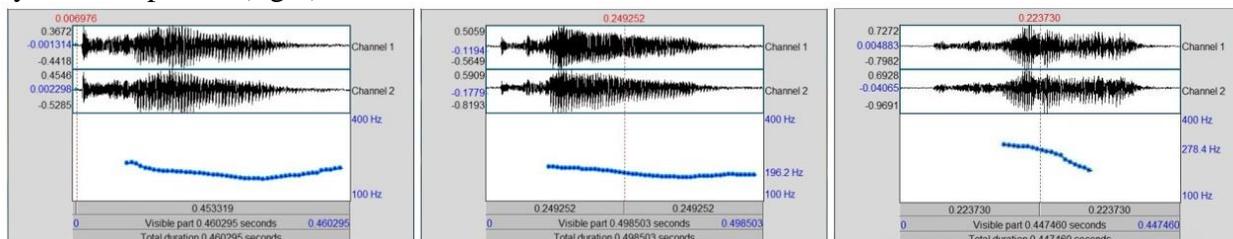
Figure 11. Pitch Curves of *kàn* ("to look") by Learner Y in Pre-test (left), Post-test (middle) and by Native Speaker (right).

DISCUSSION

The analysis of a representative sampling of the tone production data (19 words for each of the 16 participants, for a total of 304 samples) indicated, first of all, a marked ceiling effect in that 83.39% of the word tones were pronounced correctly or mostly correctly in the pre-test. This is likely due to the fact that these Chinese learners had been studying Mandarin for six months at the university level and that the selected words were all basic vocabulary items that learners were very familiar with.

In analyzing the production of individual tones, the data showed that 1st tones were pronounced correctly more often than any of the other tones, and that 4th tones were pronounced incorrectly more often than any of the other tones, corroborating earlier research by Shen (1989), who also found that 4th tones were most problematic for American learners of Mandarin.

Due to the ceiling effect and to the short duration of the training sessions, it is not possible to determine whether seeing the pitch curves had a direct effect on the learners' improvement from pre- to post-test, as found in other studies (Hardison, 2004). But the two types of comparison of pre- and post-test data both revealed some measure of improvement.

In the first rating method, raters scored randomized words from the pre- and post-tests, and a comparison of these ratings (averaged among all students and all raters) showed a marginal 1.84% improvement between pre- and post-test, with the greatest improvement of 5.42% in the most problematic 4th tone. However, a re-analysis of only the 16.61% of the tones that were incorrect in the pre-test showed that nearly half of them (49.50%) were pronounced correctly in the post-test. *T*-tests also confirmed that learners made statistically significant improvement in the incorrect tones between pre- and post-tests.

Using a second method to determine whether tone production improved, raters were presented with pre- and post-test recordings at the same time. A 17.01% improvement rate was found. In addition, a post-study survey indicated that the participants felt that seeing the pitch curves of the native speaker helped them improve their word tones (Figure 12) and that seeing their own pitch curves and practicing their pronunciation was also helpful (Figure 13).
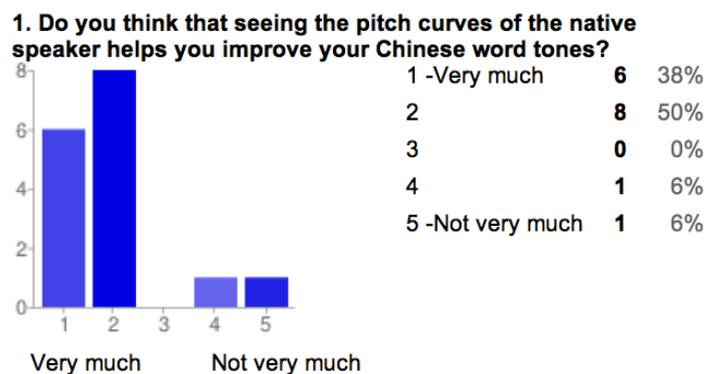
Figure 12. Student Survey Results on Helpfulness of Native Speaker Pitch Curves.

4. How helpful is it to see your own pitch curves and practice your pronunciation?

1 -Very helpful	8	50%
2	6	38%
3	1	6%
4	1	6%
5 -Not very helpful	0	0%
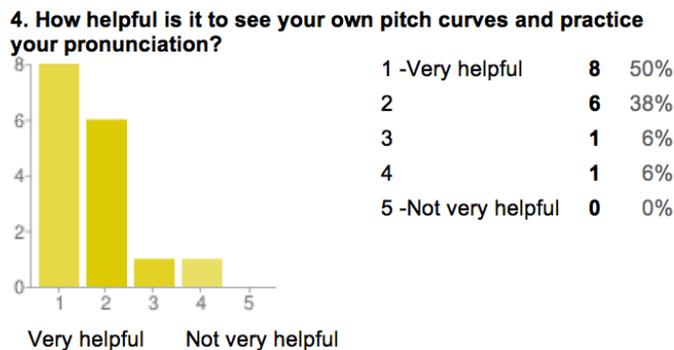
Figure 13. Student Survey Results on Helpfulness of Seeing Own Pitch Curves.

CONCLUSIONS

The purpose of this pilot study was to assess the usefulness of providing L2 Mandarin Chinese learners with visualizations of native speakers' pitch curves and the learners' own pitch curves while they tried to master Mandarin tones. The 16 learners were nearing the end of their first year of studying Chinese at the university level and proved to be very proficient at pronouncing the 19 basic, commonly used words and phrases, producing correct tones 83.39% of the time. Comparison of the learners' performance on pre- and post-tests did show improvement, with the amount of improvement varying, depending on the method of comparison. There was a statistically significant improvement in the 16.61% of the tones that were incorrect in the pre-test. The learners stated in a post-survey that they found the visualizations of the pitch curves to be helpful.

The results of this study will help in the design of a larger study with more students and not just volunteer participants. A longer, more systematic training program (i.e., more than two training sessions) will be implemented, and a more comprehensive list of words and phrases will be used, including words that might be less familiar to the learners. In addition, testing novice students who have not had as much exposure to Chinese may help reduce the ceiling effect. We anticipate that longer, sustained training with more words and phrases will provide more concrete and convincing evidence of whether visualizations of native speaker and learner pitch curves are effective for mastering Mandarin Chinese tones.

ACKNOWLEDGMENTS

ABOUT THE AUTHORS

Dorothy M. Chun (Ph.D., UC Berkeley) is Professor of Applied Linguistics and Education at the University of California, Santa Barbara. Her research areas include: L2 phonology and intonation, L2 reading and vocabulary acquisition, computer-assisted language learning (CALL) and telecollaboration for intercultural learning. She has conducted studies on cognitive process in learning with multimedia and has authored CD-ROMs and websites for language and culture acquisition. Her publications include *Discourse Intonation in L2: From Theory and Research to*

Practice, 2002, John Benjamins and articles in *CALICO Journal, CALL, Foreign Language Annals, Journal of Educational Psychology, Modern Language Journal, System,* and the online journal *Language Learning and Technology*, which she has edited since 2000.

Yan Jiang is a Ph.D. student of Gevirtz Graduate School of Education at the University of California, Santa Barbara. Her research interests are: L2 teaching and learning and computer-assisted language learning (CALL). She received a B.A. in English from Tsinghua University, China, and a M.A. in Education from San Diego State University.

Natalia Ávila is a Ph.D. student of The Gevirtz Graduate School of Education at the University of California, Santa Barbara. She received a BA in Hispanic Language and Literature and in Education, and a MA in Linguistics from Pontifical Catholic University of Chile. She has authored an online composition course in Chile. Her research interests include composition, second language writing and digital literacies.

Dorothy Chun, Department of Education, UCSB, dchun@education.ucsb.edu

Yan Jiang, Department of Education, UCSB, yan_jiang@umail.ucsb.edu

Natalia Ávila Reyes, Department of Education, UCSB, naavila@umail.ucsb.edu

REFERENCES

Chen, Q-H. (1997). Toward a sequential approach for tonal error analysis. *Journal of the Chinese Language Teachers' Association, 32*(1), 21-39.

Chun, D. M. (1998). Signal analysis software for teaching discourse intonation. *Language Learning & Technology, 2,* 61-77.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology, 8*, 34-52.

Levis, J., & Pickering, L. (2004). Teaching intonation in discourse using speech visualization technology. *System 32*(4), 505-524.

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of the Chinese Language Teachers' Association, 24*, 49-65.

Molholt, G., & Hwu, F. (2008). Visualization of speech patterns for language learning. In V. M. Holland & F. P. Fisher (Eds.), *The path of speech technologies in computer assisted language learning* (pp. 91-122). NY: Routledge.

Shen, X-N. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers' Association, 24*(3), 27-47.

Wang, Y., Sereno, J., and Jongman, A. (2006). Second language acquisition and processing of Mandarin tone. In Li, P., Tan, L.H., Bates, E., and Tzeng, O.J.L. (Eds.), *Handbook of East Asian Psycholinguistics* (Vol. 1: Chinese). Cambridge, UK: Cambridge University Press

Appendix A.

Student Demographic Questionnaire

Name:

Age:

Gender:

Year (freshman, sophomore, etc.)

Major (and Minor, if applicable):

Native/Home language:

Parents' native language:

Amount of time spent abroad and location:

Length of time studying Mandarin:

Do you speak Mandarin outside of class? With whom and how often?

Do you listen to Chinese songs or watch Chinese TV shows or movies? How often?

Other languages you speak and/or write:

IMITATION/SELF-IMITATION IN COMPUTER-ASSISTED PROSODY TRAINING FOR CHINESE LEARNERS OF L2 ITALIAN

Anna De Meo, University of Naples "L'Orientale"
Marilisa Vitale, University of Naples "L'Orientale"
Massimo Pettorino, University of Naples "L'Orientale"
Francesco Cutugno, University of Naples "Federico II"
Antonio Origlia, University of Naples "Federico II"

Recent studies on L2 acquisition, speech synthesis and automatic identification of foreign accents argue for a major role of prosody in the perception of non-native speech. Research on the relationship between pronunciation improvement and student/teachers' voice similarities has also shown that the better the match between the learners' and native speakers' voices in terms of f_0 and articulation rate, the more positive the impact on pronunciation training. This study investigates the effects of imitation and self-imitation on the acquisition of L2 suprasegmental patterns. Degree of foreign accent, improvements in intelligibility, and effectiveness of communication were measured to determine the success of each technique. For this purpose, a prosodic transplantation technique and a computer-assisted learning methodology were used.

The study was conducted with 26 Chinese speakers of L2 Italian. The stimuli used for the pronunciation training consisted of four speech acts (granting, order, request and threat) uttered by the 26 Chinese and two Italian native speakers, and the items necessary for the self-imitation training were obtained through prosodic transplantation, i.e. transferring suprasegmental features from native speakers' voices to the L2 ones. Chinese students divided into two different groups practiced imitation and self-imitation and the self-imitation impact was evaluated by comparing pre- and post-training performances of both groups. Both teaching strategies promoted an improvement in learners' performances; however, the self-imitation training proved to result in more accurate prosodic realizations.

INTRODUCTION

Pronunciation improvement is a relevant issue in the area of spoken language technology for language learning (Chun, 2013; Eskenazi, 2009; Levis, 2007; Martin, 2012). However, during the last two decades, only a few studies, carried out on learners with different L1s and focusing on different target languages, have investigated the relationship between the student/teacher voice similarity and pronunciation improvement (Bissiri, Pfitzinger, & Tilmann, 2006; Jilka & Möhler, 1998; Nagano & Ozawa, 1990; Peabody & Seneff, 2006; Sundström, 1998; Tang, Wang & Seneff, 2001). Results from these studies have shown that the better the match between the learners' and native speakers' voices in terms of f_0 and articulation rate, the more positive the impact on pronunciation achievement, suggesting the existence of a user-dependent golden speaker (Probst, Ke, & Eskenazi, 2002). As claimed by Felps, Bortfeld, & Guiterrez-Osuna (2009), it would be beneficial for L2 students to be able to listen to their own voices producing utterances in a native accent. As a consequence, the most effective golden speaker to learn

segmental and suprasegmental features of a second language is the learner's own voice with a native accent.

This study investigates the effects of the self-imitation strategy, i.e. the speaker's imitation of his/her own voice properly modified according to the target native model on the acquisition of L2 suprasegmental patterns, comparing the results with those achieved with traditional imitation exercises. For this purpose, a prosodic transplantation technique and a computer-assisted learning methodology were used. To determine the success of each technique, four different variables were considered: speech act identification, communication effectiveness, degree of foreign accent and improvements in intelligibility.

MATERIALS AND METHOD

Technique

This study is based on the use of the rhythmic-prosodic transplantation technique (Pettorino & Vitale, 2012; Yoon, 2007), which makes it possible to transfer one or more acoustic parameters (pitch, intensity, articulation rate, frequency and duration of silent pauses) from a native speaker (the "donor") to a non-native speaker (the "receiver"), without altering the segmental sequence and the identity of the synthesized voice. This technique is based on the PSOLA (Pitch-Synchronous Overlap and Add) algorithm (Charpentier & Moulines, 1989), implemented in Praat (Boersma, 2001).

The transplantation procedure involves a fixed sequence of steps, divided into five phases: treatment of anomalies, segmentation, transplantation of duration, transplantation of intensity, and pitch contour superimposition. All these operations have been automatized through a Praat script and then applied to the voices selected for this study.

Stimuli

For the purpose of this research, two Italian sentences were chosen. The meaning of these two sentences can vary by using different pitch contours, even if the syntactic structure is kept unchanged. Human languages, indeed, generally allow the speaker to express the modal meaning of the sentence, i.e., attitude towards the message content, by using different pitch contours (Sorianello, 2006). For example, a sentence such as "We were sharing a hamburger" can be uttered and interpreted as a question or a statement, depending on the different modulation of the fundamental frequency (f_0) contours.

For this study, four speech acts were considered (*request, order, granting* and *threat*). The first three were already tested in previous studies (De Meo & Pettorino 2011; De Meo, Pettorino & Vitale, 2012, Pettorino, De Meo & Vitale, 2012), while *threat* was newly introduced in order to create a balanced set of stimuli, composed of the same number of pragmatically basic and complex utterances. In fact, the four considered speech acts are characterized by different degrees of familiarity for NNSs. Requests and orders are always introduced early in the learning process and, at the same time, are more frequent in the language input. Grantings and threats, by contrast, are rarely presented in advanced level language courses.

The sentences and the different meanings considered are the following:

> Sentence 1 - "Lascia i piatti sul tavolo"
> > granting ("Ok, you can leave the dishes on the table.") (Figure 1)
> > order ("Leave the dishes on the table!") (Figure 2)

Sentence 2 - "Ne parliamo stasera a casa"
          request ("Shall we talk about it at home tonight?") (Figure 3)
          threat ("We will talk about it at home tonight.") (Figure 4)
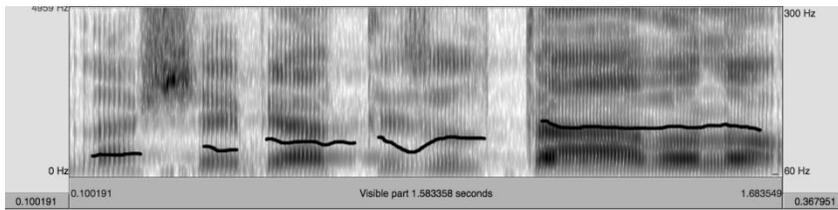
Figure 1. Granting, male voice.
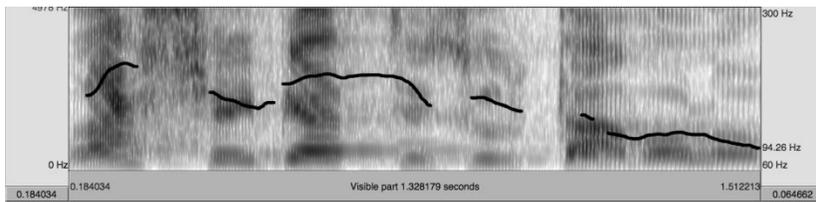
Figure 2. Order, male voice.
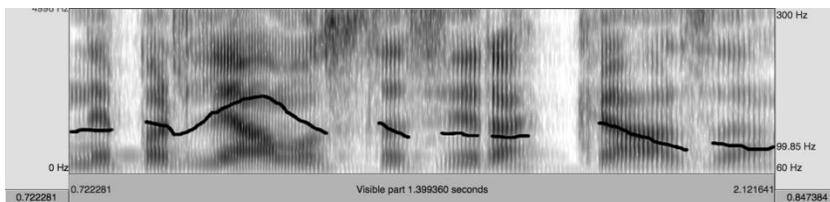
Figure 3. Request, male voice.
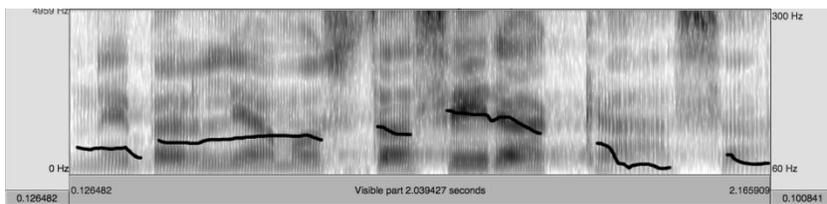
Figure 4. Threat, male voice.

The four sentences were recorded by the NSs and the NNSs involved in this study in an anechoic chamber. With the help of a professional translator, it was ensured that the non-native subjects had understood the true meaning associated with each speech act before the recording phase.

Subjects

Speakers

The subjects involved in this study were 6 native Italians (3 males and 3 females) and 30 L1 Chinese speakers (25 females and 5 males), all having a high-intermediate level of competence in Italian (B2 of the Common European Framework of Reference). Both groups were university students living in the Campania Region, Southern Italy.

Listeners

Two different groups of native Italian listeners were involved in this study: 30 subjects (Group 1) and 52 subjects (Group 2). All the listeners, male and female with an average age of 24, lived in

Naples and were accustomed to the same diatopic variety of Italian as the six native Italian speakers.

PROCEDURE

The entire experimental procedure can be divided into five main steps:

1. two pre-tests to select the native speakers (NSs) to be involved in the study and to confirm the presence, already detected by researchers, of a strong foreign accent in the non-native voices;

2. a perceptual test to select the non-native speakers (NNSs) to be trained;

3. a rhythmic-prosodic transplantation of the NNSs' utterances;

4. imitation and self-imitation prosodic training;

5. a final perceptual test to compare imitation and self-imitation post-training performances.

Step 1

In order to select the most suitable native and non-native subjects for this study, two different pre-tests were carried out.

Since the prosodic model to be used for both the imitation and the self-imitation trainings was offered by the native Italian speakers, it was important to choose the most communicatively accurate native male and female voices. Furthermore, we decided to select only Chinese learners having a strongly accented L2 speech, in order to easily observe the post-training effects.
To this end, the four utterances for the stimuli were recorded by the six NSs and administered in random order to the listeners of Group 1, who were asked to:

- identify the speech act (multiple choice task);
- evaluate the communication effectiveness on a five-point-scale (1=min, 5= max);
- evaluate the degree of foreign accent on a three-point-scale (native accent, mild foreign accent, strong foreign accent);
- assess intelligibility on a three-point-scale (poor, sufficient, good).

Regarding the Chinese voices, all NNSs were instructed to read a short text in Italian, then assessed by the listeners of Group 1 in a second test session. Hence, for each non-native voice the degree of foreign accent (native accent; mild foreign accent; strong foreign accent) was evaluated.

As far as the L1 Italian subjects are concerned, the pre-test was used to choose the best male and female performances for each of the four provided utterances. As was to be expected, although all the NSs received a positive evaluation (the average correct speech act identification was of 72%), two of them appeared to be more communicatively accurate than the others (correct recognition: male voice 82%; female voice 90%). These two speakers were the only voices to be finally involved in the study, since there were no other significant differences among all the NSs, in terms of both degree of foreign accent and intelligibility.

With respect to the second pre-test, results showed that 26 out of 30 NNSs were judged as prevalently strongly foreign accented (73%). The remaining 4 NNSs were excluded from the experiment.

Step 2

As the 26 Chinese learners had a high-intermediate level of competence in Italian, they were already able to produce some acceptable utterances. It was important to exclude from the training sessions utterances that were properly uttered and therefore perfectly acceptable from the start. This was accomplished by administering a perceptual test to Group 2 listeners.

For this second step, the four sentences recorded by the 26 NNSs and by the 2 NSs were randomly arranged. The L1 Italian utterances were used as control elements to set the threshold of acceptability for the L2 productions. A total of 112 stimuli (28 x 4) was administered in a perceptual test, specifically devoted to the evaluation of all the involved subjects on each of the following aspects:

- speech act identification (multiple choice task);
- communication effectiveness (five-point scale rating: 1=min, 5=max)
- intelligibility (three-point scale assessment: poor, sufficient, good).

The minimum percentage of correct speech act identification obtained by the NSs, which was approximately 60%, represented the threshold for the selection of the NNSs productions to be accepted for the training phase. In other words, all the speech acts that gained a percentage of correct identification higher than 60% were considered acceptable and not in need of training. On the basis of the set threshold, only 68 NNSs' utterances were selected (26 grantings, 21 threats, 10 orders and 11 requests). Figure 5 shows the mean percentage values of correct speech act identification concerning only the 8 NSs' and the 68 NNSs' utterances selected for the training.
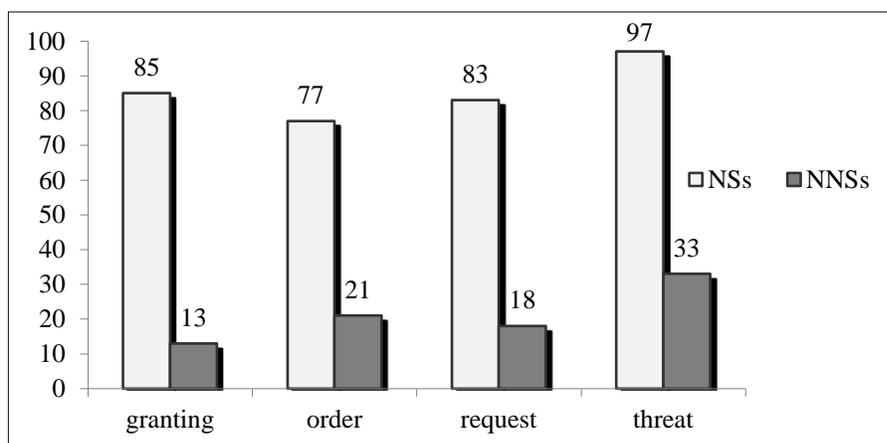
Figure 5. Correct speech act identification (mean percentage values). NSs and NNSs to be trained in comparison.

It should be noted that for the 55% of the errors produced in the speech acts recognition assessment, the judging audience gave "order" as their answer, probably because the prevailing flat pitch of the Chinese L2 Italian speech coincides with the pitch profile of the order in L1 Italian. The remaining 45% of the errors were distributed fairly randomly among the other speech acts (granting 14%, request 19%, threat 12%).

The gap existing between native and non-native speakers' achievements is evident also in terms of communication effectiveness, since Italians were given an average score of 4.7, while the Chinese learners only reached a 2.5 level. No clear variations were observed between the

different speech acts (Figure 6). As for intelligibility, the NSs, as expected, were fully understood by all the listeners ("good" intelligibility: 98%), while the NNSs, although they also got 61% "good" evaluations, were judged as just sufficiently intelligible by the 33% of the evaluators and not understandable at all by the remaining 6%.
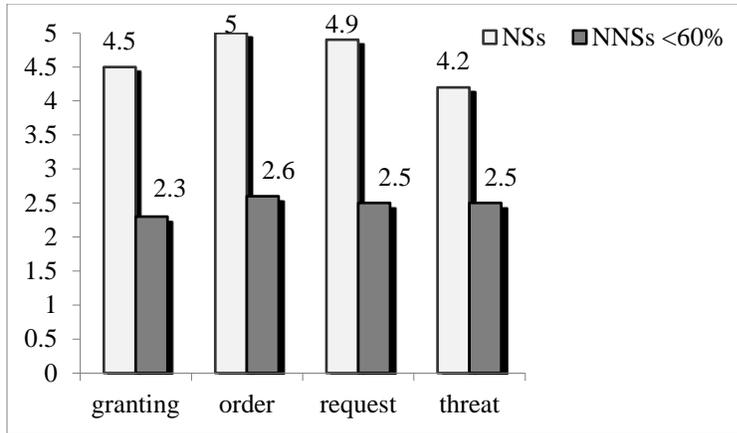
Figure 6. Communication effectiveness per speech act of the NSs and the selected NNSs (1=min, 5=max).

Step 3

All the NNSs correctly produced at least one of the proposed speech acts. Of the 68 NNSs' items finally selected, half underwent the imitation treatment and the other half the self-imitation one. Although a partially random subdivision was carried out, an attempt was made to maintain a balanced distribution of the two treatments in terms of speech acts and speakers. The prosodic transplantation was performed only on the 34 items to be used for the self-imitation training.
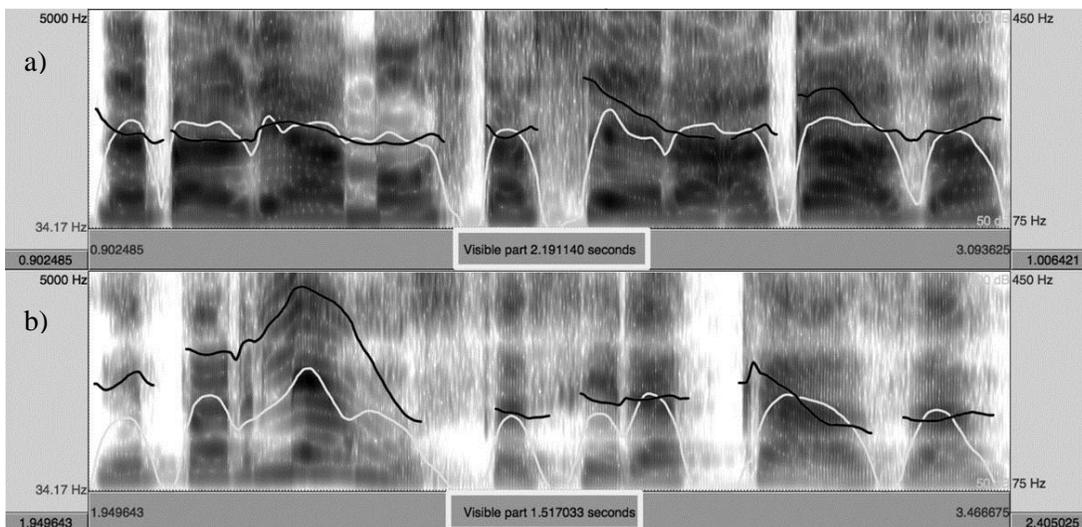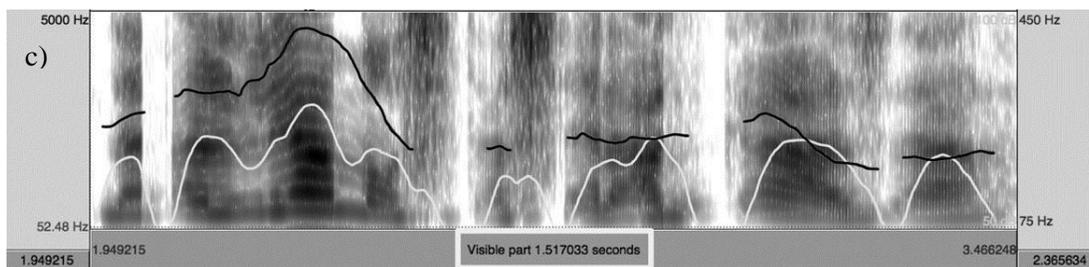
Figure 7. Request "Ne parliamo stasera a casa?", female voices. a) Chinese original voice; b) Italian model voice; c) Chinese synthesized voice. Black line: pitch contour – White line: intensity.

As can be seen in Figure 7, the Chinese original voice (Figure 7a) receives the same segmental duration, intensity contour and pitch movement of the Italian "donor's" voice (7b) resulting from the transferring of the rhythmic-prosodic features (7c). With the exception of micro variations that are not perceptually relevant, the temporal extension of the whole sentence (as indicated in the box "visible part" below each spectrogram) and, more specifically, the duration of each segment of the manipulated utterance are comparable to those of the Italian model voice. With regards to the pitch contour, as a result of the transplantation procedure, the flattened movement of the original Chinese voice (7a) was transformed into a much more varied intonation, with a marked peak on the third syllable and slighter pitch shifts on the second part of the utterance (Figure 7c). As shown by the continuous grey line on the spectrogram, with the transplantation even the energy peaks are perfectly repositioned according to the model voice.

Step 4

The audio files to be used for the imitation and self-imitation trainings were arranged into 26 exercise packages, each containing a different series of audio files to be used by one of the 26 NNSs. Each NNS was asked to listen to the utterances produced by a native speaker's voice or by his/her own synthesized voice contained in his/her own package and to exercise in imitating the input. After 5 minutes of training, learners were instructed to record the output sentences.

Step 5

The collected post-training performances were used to arrange a final perceptual test to compare the improvement induced by the two treatments. To this end, both the pre- and post- training productions were assembled in random order and administered to the Group 2 listeners, who were asked to perform the same task as in the Step 2 perceptual test. In this case the foreign accent assessment was added (three-point scale evaluation: native accent, mild foreign accent, strong foreign accent) in order to evaluate the impact of the exercises that were carried out.

RESULTS

As it can be seen in Table 1, both trainings improved the NNSs' rhythmic-prosodic performances, although the results obtained by the self-imitation exercises seem to be more relevant for the order and the request. Results are statistically significant ($p < 0.001$, ANOVA).

Table 1
Speech Act Identification Improvement (Δ values: post training % – pre training %)

	Imitation	Self-imitation
Granting	+47%	+50%
Order	+22%	+48%
Request	+47%	+71%
Threat	+26%	+33%

Communication effectiveness also underwent a slight general improvement, but even in this case the self-imitation treatment was more effective, especially for the granting and the request functions (Table 2). Results are statistically significant ($p < 0.001$, ANOVA).

Table 2
Communication Effectiveness Improvement (five-point scale evaluation, Δ values: post training – pre training)

	Imitation	Self-imitation
Granting	+0.9	+1.3
Order	+0.5	+0.3
Request	+0.3	+1.5
Threat	+0.3	+0.4

The two trainings under investigation produced very similar improvements in terms of intelligibility (Table 3), while in the foreign accent reduction the results obtained by the self-imitation are slightly more positive: the post-imitation utterances have produced a decrease of the "strong foreign accent" in favour of the "mild foreign accent", whereas the post self-imitation productions were judged as being of a native accent by a small percentage of the listeners (Table 4). Even in this case results prove to be statistically significant ($p < 0.001$, ANOVA).

Table 3
Intelligibility Improvement (Δ values: post training % – pre training %)

	Imitation	Self-imitation
Poor	+4%	+4%
Sufficient	-21%	-22%
Good	+17%	+19%

Table 4
Foreign Accent Improvement (Δ values: post training % – pre training %)

	Imitation	Self-imitation
Native accent	0	+6%
Mild foreign accent	+27%	+19%
Strong foreign accent	-27%	-25%

CONCLUSIONS

This study shows that computer-assisted prosody training based on both imitation and self-imitation produces good results in terms of pronunciation improvement, providing a spin-off for

prosody learning, communication effectiveness and intelligibility improvement, and foreign accent reduction. However, self-imitation, made possible by the use of the rhythmic-prosodic transplantation technique, generally achieves more satisfactory results.

These results led us to develop a project that aims at creating a software program, ProsoTrainer, devoted to the prosodic pronunciation improvement for learners of L2 Italian, favouring at the same time foreign accent reduction and the improvement of communication effectiveness. By means of the prosodic-intonational transplantation procedure, the suprasegmental features of the native speaker (pitch, intensity, articulation and speech rate, frequency and duration of pauses) would be cloned and transferred in real time to the L2 learner's voice, without altering the perception of the L2 speaker's identity. The learner's voice thus becomes the "native" model to imitate. However, speakers with different L1s and different levels of L2 competence, and a greater number of speech acts have to be considered in order to get sufficient data to support the development of a technological tool that makes teaching and/or autonomous learning of the L2 suprasegmental features easier.

ABOUT THE AUTHORS

Anna De Meo is Professor at the University of Naples "L'Orientale," Italy. She teaches Multiculturality and Language Acquisition and Specialized Translation. Her main fields of interest are L2 acquisition (with special reference to Italian), L2 speech perception and production, interlanguage pragmatics, and translation. Email: ademeo@unior.it +39-0816909031 University of Naples "L'Orientale" - via Nuova Marina 59 - 80133 Naples - Italy

Marilisa Vitale is a PhD student in Linguistics at the University of Naples "L'Orientale," Italy, and currently trainee at the LIMSI-TLP (Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur - Groupe Traitement du Langage Parlé) of the CNRS in Paris. Her main fields of interest are rhythmic-prosodic aspects of L2 Italian, foreign accent and credibility correlates, speech manipulation, and teaching L2 Italian prosody. Email: marilisavitale@hotmail.it +39-0816909031 University of Naples "L'Orientale" - via Nuova Marina 59 - 80133 Naples - Italy

Massimo Pettorino is Professor of Experimental Phonetics at the University of Naples "L'Orientale," Italy, and Director of the GSCP, a special interest group on spoken communication of the Italian Linguistic Society (SLI). He works in the field of acoustic and prosodic analysis of speech. Email: mpettorino@unior.it +39-0816909912 University of Naples "L'Orientale" - via Duomo 219 - 80139 Naples - Italy

Francesco Cutugno is Assistant Professor at the University of Naples "Federico II." His main research interests concern speech technologies and information retrieval in spoken language corpora. He is also interested in automatic prosodic analysis, syllable segmentation and rhythm. He is responsible for the Language Understanding and Speech Interfaces (LUSI) Lab in the Department of Physics. Email: cutugno@unina.it +39-081676850 Department of Physics – University of Naples "Federico II" - via Cinthia - 80126 Naples - Italy

Antonio Origlia obtained a Master degree in Computer Science at the University of Naples, "Federico II" in 2009 and is currently undertaking a PhD in mathematical and computational sciences at the Language Understanding and Speech Interfaces (LUSI) Lab. His area of interest mainly covers affective computing with focus on emotional speech. He is also interested in prosodic analysis tools development, automatic syllable segmentation, and prominence detection

and rhythm. Email: antonio.origlia@unina.it +39-081679961 Department of Physics – University of Naples "Federico II" - via Cinthia - 80126 Naples – Italy

REFERENCES

Bissiri, M. P., Pfitzinger, H.R., & Tillmann, H.G. (2006). Lexical stress training of German compounds for Italian speakers by means of resynthesis and emphasis. In *Proceedings of the 11th Australian International Conference on Speech Science & Technology* (pp. 24–29). University of Auckland, New Zealand.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International 5*, 341-345.

Charpentier, F., & Moulines, E. (1989). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. In *Proceedings of the First European Conference on Speech Communication and Technology - Eurospeech* (pp. 2013-2019). Paris: European Speech Communication Association.

Chun, D.M. (2013). Computer-assisted pronunciation teaching. In C.A. Chapelle (ed.) *The Encyclopedia of Applied Linguistics* (pp. 823-834). Oxford, UK: Wiley-Blackwell.

De Meo, A. & Pettorino, M. (2011). L'acquisizione della competenza prosodica in italiano L2 da parte di studenti sinofoni. In E. Bonvino & S. Rastelli (eds.), *La didattica dell'italiano a studenti cinesi e il progetto Marco Polo*. Proceedings of the 15th Workshop AICLU (pp. 67-78). Pavia: Pavia University Press.

De Meo, A., Pettorino, M., & Vitale, M. (2012). Comunicare in una lingua seconda. Il ruolo dell'intonazione nella percezione dell'interlingua di apprendenti cinesi di italiano. In M. Falcone & A. Paoloni (eds.), *La voce nelle applicazioni*. Proceedings of the 7th Congress of Italian Association of Speech Sciences AISV (pp. 117-129). Roma: Bulzoni editore.

Eskenazi, M. (2009). An overview of spoken language technology for education. *Speech Communication, 51*(10), 832-844.

Felps, D., Bortfeld, H., & Gutierrez-Osuna, R. (2009). Foreign accent conversion in computer assisted pronunciation training. *Speech Communication, 51*(10), 920-932.

Jilka, M., & Möhler, G. (1998). Intonational foreign accent: speech technology and foreign language teaching. In P*roceedings of ESCA Workshop on Speech Technology in Language Learning* (pp. 115–118).

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27*, 184-202.

Martin, P. (2012). Automatic prosodic comparison between model and imitation sentences in a second language teaching computerized environment. In A. De Meo & M. Pettorino (eds.), *Prosodic and rhythmic aspects of L2 acquisition. The case of Italian* (pp. 263-276). Newcastle upon Tyne: Cambridge Scholars Publishing.

Nagano, K., & Ozawa, K. (1990). English speech training using voice conversion. In *1st International Conference on Spoken Language Processing* (ICSLP 90) (pp. 1169-1172). Kobe, Japan.

Peabody, M., & Seneff, S. (2006). Towards automatic tone correction in non-native mandarin. In *Proceedings of the 5th international conference on Chinese spoken language processing* (pp. 602-603). Kent Ridge, Singapore.

Pettorino, M., De Meo, A., & Vitale, M. (2012). La competenza prosodico-intonativa nell'italiano L2. Analisi e sintesi del segnale fonico di cinesi, vietnamiti e giapponesi. In S. Ferreri (ed.), *La linguistica educativa. Proceedings of the 44th International Congress of Italian Linguistics Society SLI* (pp. 329-342). Roma: Bulzoni.

Pettorino, M., & Vitale, M. (2012). Transplanting native prosody into second language speech. In M.G. Busà Maria Grazia, A. Stella (eds.), *Methodological Perspectives on Second Language Prosody. Papers from ML2P 2012* (pp. 11-16). Padova: CLEUP.

Probst, K., Ke, Y., & Eskenazi, M. (2002). Enhancing foreign language tutors - in search of the golden speaker. *Speech Communication 37*(3-4), 161-173.

Sorianello, P. (2006). *Prosodia*. Roma: Carocci.

Sundström, A. (1998). Automatic prosody modification as a means for foreign language pronunciation training. In *Proceedings of ISCA Workshop on Speech Technology in Language Learning* (STILL 98) (pp. 49-52). Marholmen, Sweden.

Tang, M., Wang, C., & Seneff, S. (2001). Voice transformations: From speech synthesis to mammalian vocalizations. In *Proceedings of Eurospeech 2001*, Aalborg, Denmark.

Yoon, K. (2007). Imposing native speakers' prosody on non-native speakers' utterances: The technique of cloning prosody. *Journal of the Modern British & American Language & Literature 25*(4), 197-215.

IMPROVING ORAL PROFICIENCY BY RAISING METACOGNITIVE AWARENESS WITH RECORDINGS

Jessica S. Miller, University of Wisconsin-Eau Claire

Audio technology such as podcasting may be an efficient tool to improve oral proficiency by exposing learners to more authentic input (e.g., listening to podcasts) and by giving them opportunities to produce substantially more output (e.g., creating podcasts). Combined with self-evaluation activities, voice recording technology can also contribute to metacognitive awareness, i.e., "the ability to reflect upon, understand and control one's learning," (Schraw and Dennison, 1994), an important component of the learning process (Swain, 1985).

This exploratory study examines multiple learning activities in conversation courses designed to develop spontaneous second language (L2) oral proficiency. It also investigates learners' perception of what is needed to improve. Data were gathered in two university L2 conversation courses: an intermediate German course, and an advanced French course. Some of the activities proposed in those courses employed voice recording technology and rubrics to enhance metacognition skills, while others did not.

Preliminary results show that students gained from using audio technology and rubrics as they reported benefitting from paying more attention to their speech, which helped them correct their errors. Although not all favored those exercises, a vast majority acknowledged the need to be pushed in order to become proficient, no matter the activity in which they engage. Developing activities that enhance their Willingness To Communicate (McCroskey & Richmond, 1991; MacIntyre 2007) as well as their metacognitive awareness seemed to have a positive impact.

INTRODUCTION

Short of traveling with a class to practice the target language in an authentic setting, instructors can bring the target culture to the classroom via technology. Because technology is now a part of students' everyday lives, using it to teach and learn foreign languages makes the learning experience more relevant and applicable to their lives. For example, some authors have highlighted how podcasts can help second language (L2) students. Chinnery (2006) noted the advantage of portability, which can increase accessibility to L2 input. Ducate & Lomicka (2009) found that students appreciated the creative aspect of producing their podcasts. So did O'Bryan & Hegelheimer (2007) who showed that learners viewed podcasting in a positive way. Lord's research (2008) demonstrated that students' attitudes and pronunciation abilities in Spanish improved as they engaged in crafting their podcasts. Onsrud's high school French project (2009) concluded that students who recorded their assignments acquired oral proficiency skills more successfully.

Such positive reports could be explained by the fact that when students produce and listen to podcasts in another language, they create "their own mobile immersion environments" (Thorne & Payne, 2005, p. 386). They can in this manner immerse themselves in authentic material on topics of their own choosing, which makes that activity attractive (Rosell-Aguilar, 2007), contrary to traditional language lab settings where one listens and speaks around theme generally imposed by a textbook. For example, students in Miller & Hlas (2010) made a series of three

podcast episodes in their intermediate French and Spanish courses. Each team of two selected an overarching theme for all episodes such as healthy foods, traveling, university life, etc. Every installment had a different structure within which students had room for creativity, but all were linked by the general theme and together formed a radio show: episode one was an interview, episode two a top-ten list, and episode three a debate. This semi-guided activity allowed students to use their target language creatively in and outside of class as they researched their themes to prepare the radio program. The language lab then became a mock radio station where shows were produced and broadcasted to peers, thus creating a mini-immersion experience. The content learned from listening to and creating personalized podcasts could later be shared with classmates and recycled for further speaking practice.

In addition, technology that allows learners to monitor and evaluate their own language likely promotes metacognitive skills, an aspect of L2 acquisition that, if strengthened, supports language learning (Anderson, 2008; Graham, 2006; O'Malley & Chamot 1990; Rivers, 2001; Thompson, 2012). Metacognition can be defined as "the ability to reflect upon, understand and control one's learning," (Schraw & Dennison, 1994). In other words, developing students' metacognitive skills is helping them think about thinking so they develop the capacity to evaluate themselves and control their own learning. Fostering metacognitive skills can be done by asking learners what they find difficult and why (Tanner, 2012), or by addressing study habits (Schraw & Dennison, 1994), or by modeling instructors' or peers' thought processes when dealing with difficulties (Schraw, 1998).

Developing metacognition in L2 learners is essential for attention to be drawn to correct language uses. Swain's Output Hypothesis (Swain, 1985) indicates that processes such as noticing and metalinguistic reflection facilitate second language acquisition. Indeed, research has shown that across any field learners with higher metacognitive skills pay closer attention and use more effective learning strategies (Cooper & Sandi-Urena, 2009), and as a result "perform better than unaware learners" (Schraw & Dennison, 1994). Active learning paired with metacognition development contributes to greater learning in that students process events better and consequently memorize them more precisely. They also remember the reasons why they were asked to engage in those activities, as well as what they learned and how they learned from them (Tanner, 2012).

The present exploratory study investigates learners' perception of what is needed to improve L2 oral proficiency as it relates to metacognitive skill development assisted by recording technology. Data were collected from two university L2 conversation courses that integrated recording technology combined with self- and peer-monitoring rubrics. In a 300-level German conversation class, students recorded multiple narrations of stories using scripted and unscripted speech. In a 400-level French conversation class, students produced podcasts as well, following Miller & Hlas' methods (2010). In pairs they created three episodes of their own radio show as described above. In both language groups learners were required to evaluate their own recordings. In the French class only they were asked to evaluate peers' podcasts.

RESEARCH QUESTIONS

Current professional demands highlight second language proficiency as a crucial attribute for a global workforce (Henn, 2012) but budget constraints make it increasingly difficult to send students abroad or to limit class sizes for more individual attention. If used thoughtfully, technology may increase the quantity and quality of learners' input and output, and allow

learners to self-monitor and self-evaluate. They may thus develop metacognitive behaviors. For example, Lord (2008) found that self-evaluation by language learners can help them acquire proper pronunciation by raising metalinguistic awareness. This exploratory study on recording technology and oral proficiency aims at offering ways to make instruction more effective and relevant to current needs. It also reflects on changing learners' behaviors by showing them different learning strategies that have the potential to increase their autonomy and thus enhance their L2 capabilities. Foreign language college students exposed to different activities including recorded and self-evaluated tasks are asked the following broad questions: what learning activities do you feel help increase their oral proficiency? What do you feel is needed to become proficient? Within those questions responses are scrutinized to investigate whether or not learners benefit from recording technology, and if so, whether metacognition has a role to play in those benefits. Pedagogical implications are discussed as well.

METHODS

Qualitative data were collected in 2012 at a mid-sized liberal arts university in the Midwest from students enrolled in two L2 conversation courses: an intermediate 300-level German course and an advanced 400-level French course. In both, learners practiced spontaneous speech in multiple ways, with and without recording technology. They were surveyed at the end of the term.

Participants

Intermediate and advanced language courses have limited enrollment so as to ensure frequent individual interaction between learners and instructors. Among all students registered in those two classes, eight learners of French and ten learners of German chose to participate in this study. While the results cannot be generalizable due to the small sample size, they indicate possible trends for further explorations.

Procedure

For one semester all learners received instruction meant to improve their oral proficiency skills according to the national standards developed by the American Council on the Teaching of Foreign Languages (ACTFL, 2002). While both courses were taught by different instructors and used different types of activities, the course objectives and teaching methods were similar. One goal of both courses was to prepare students for an ACTFL's Oral Proficiency Interview (OPI), an oral exam required to receive teaching licensure in many states. An online survey containing open-ended questions designed to trigger reflection on oral proficiency was distributed at the end of the semester. Those questions appear in the Results section below.

Activities performed in both the German and French classes were debates, conversations, presentations, summaries of articles, and story narrations. In the French class, learners also listened to authentic podcasts, produced their own, and evaluated theirs and their classmates'. Varying degrees of attention were given to those tasks depending on the course. Table 1 below gives an overview of what types of activities were used in each class. It shows which ones used voice recordings that students could hear later, which ones did not, which ones alternated, and which ones were accompanied by a self-evaluation.

Table 1.

Learning activities designed to increase oral proficiency

Activities	French course	German course
Podcasting	recorded + self-evaluated	N/A
Debates	recorded and not recorded	not recorded
Conversations	recorded and not recorded	not recorded
Presentations	not recorded	not recorded
Summaries	not recorded	not recorded
Story narrations	recorded and not recorded + self-evaluated	recorded and not recorded + self-evaluated

RESULTS

Among the activities that you identified above, describe the one that you feel helped improve your foreign language oral proficiency the most.

Learners of German ranked story narration as the most helpful followed by article summaries (Table 2 below). A participant explained: "Telling stories using only pictures as guides really helped improve my oral proficiency because it forced me to really think about narration in the past tense, word order, and conjugations. As the semester went on, I was able to catch my mistakes more quickly and be more confident in speaking". Another one describes: "Not only did we have to narrate a story and recite without reading it, but then we also listened to our recordings, found the mistakes, and corrected them. It was really helpful listening to our own speech". Those comments underline that narrating and evaluating one's performance may support the development of oral proficiency by raising consciousness to some issues. Other participants' comments stressed the importance of developing the ability to speak spontaneously in a pressure-free environment. For example, one participant noted that summaries "forced me to be able to think quickly and put all the words together in an order that worked within the grammatical structures" while another remarked that "conversing with partners seemed to be the most effective because it allowed me to speak my second language in a less stressful situation than having the entire class listen to what I'm saying".

Table 2.

Activities that German students found helpful to increase proficiency

Rank	Type of activity	Number of students who cited it
1	Story narration	4
2	Article summaries	3
3	Role-plays	1
3	Conversations	1
3	OPI training	1

Students of French selected debates as being the most helpful (Table 3 below). One participant explained: "I think that the in-class debates were the most helpful for improving my foreign language oral proficiency. They forced us to talk about topics that we wouldn't on a daily basis. While we had time to prepare for some of them, others were spontaneous". Many comments in the French group echoed those in the German group regarding the usefulness of practicing spontaneous speech in a relaxed environment. Debates were not graded, their topics varied form silly to serious, and students could rely on their peers and the instructor for assistance. As for podcasts, one student wrote: "I feel the creation of the podcasts greatly helped improve my oral proficiency. The *table française* [conversation table outside of class] is great too, but I tend to be too shy to speak up a lot […]. However, during the podcasts, I am forced to speak an equal amount as my partner. Also, the podcasts were spontaneous, so it was great practice for spontaneous discussion of ideas that were complex and beyond everyday conversation topics". In this case, the learner benefitted from pressure linked to the obligation to speak.

Table 3.

Activities that French students found helpful to increase proficiency

Rank	Type of activity	Number of students who cited it
1	Debates	4
2	Creating podcasts	2
2	Conversations	2

Interestingly, the verb "forced to" occurs frequently in both groups' positive feedback. The lack of stakes and accountability during traditional conversations seems to negatively affect participation rate, and therefore less progress is accomplished. The students acknowledged the usefulness of being pushed to talk. Tying those responses to the Willingness to Communicate (WTC) framework (McCroskey & Richmond, 1991) can help interpret the results. The WTC concept was applied to L2 acquisition and defined within that context as "the probability of speaking when free to do so" by MacIntyre (2007, p. 564), with or without the speakers' explicit

awareness. When the participants explained the benefits of being "forced" to speak, they were in a way describing convincing external factors motivating them to communicate willingly, even though they may not have been aware of it. MacIntyre explained: "Perhaps the most important decision language learners face is whether or not to use their incipient skills when the opportunity arises, inside or outside the classroom" (p. 573). Therefore the activities described as helpful – specifically the recorded story narrations and non-recorded article summaries in the German group, and the debates (either recorded or not) and the recorded podcasts in the French group – have in common that they increased the students' WTC. Those activities not only led to more output, but to more reflection about the learning process and the errors to be corrected.

What does a speaker need to become proficient? Name at least 3 elements that, in your opinion, characterize a high level of oral proficiency.

The answers from the German group varied greatly but a majority believed that correct grammar is the most important, followed by vocabulary. Among the learners of French, the results differed. While only one respondent in the German group cited pronunciation, that was the most frequently cited component as a correlate of proficiency in the French group. Grammar came in third position, after guidance. Here, guidance as expressed by students likely means corrective feedback. For example, learners often request the presence of a native speaker to point out and correct their errors during conversation tables. This finding may reflect a certain anxiety toward pronunciation experienced by learners of French who tend to find that aspect quite challenging. The same may be true for learners of German and grammar. The German group cited "confidence" in third place, also indicating the need to let go of fears tied to insecurities in order to become proficient.
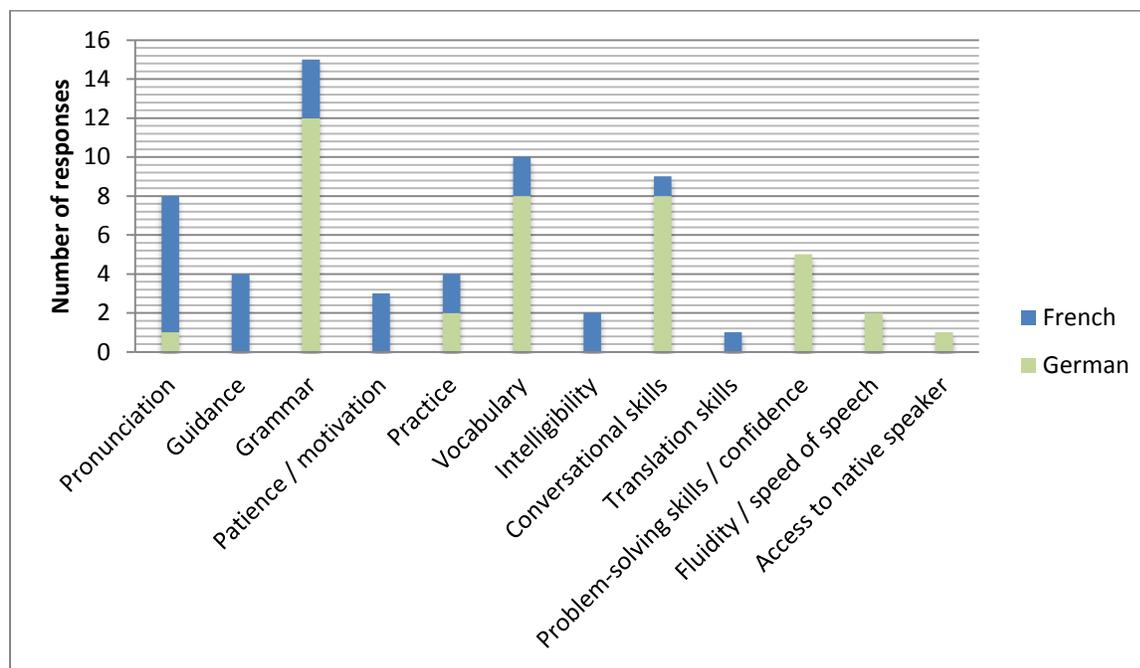
Figure 1. Elements characterizing high proficiency according to students

Among your activities outside of class, and reflecting on your entire life as a language learner, describe an event or activity that helped improve your oral proficiency the most.

The most frequent answer to this question was "immersion" for both groups. For the learners of German, equally as important were self-monitoring through recordings, receiving OPI training, and communicating with natives outside of immersion contexts. For instance, a student of German noted: "It's good to be corrected/correct your own speech soon after you've said it so that you learn from your mistakes and hopefully quit making them". That comment illustrates that outside of immersion, being made to notice mistakes is perhaps one of the better teaching and learning techniques.
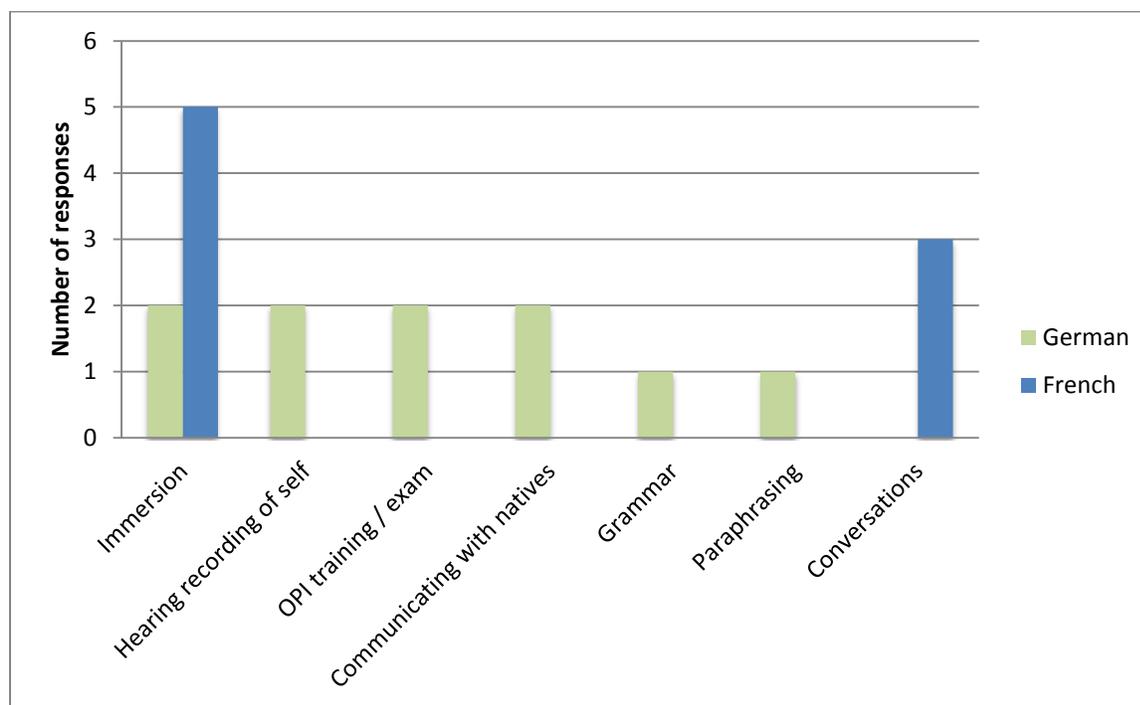
Figure 2. Elements described by students as helpful to increase oral proficiency

In the French group, the majority responded that immersion was important: "The most transformative experience was studying abroad in France for a semester. It greatly helped that I had more than 4 hours of class everyday only in French. I also lived with a host family who only spoke French. This mainly improved my oral proficiency out of necessity". Otherwise they cited conversations, noting that immersion and conversations display similarities: "Definitely any immersion scenario has helped me the most. In French classes we have discussion groups that are mini immersion experiences and they've helped me out a lot. When I am forced to speak the language in order to get by, it helps me go outside my comfort zone and it helped me become much more confortable with the language". Once again, words like "necessity" and "forced to" emerged. There is an apparent contradiction between learners needing to be pushed outside of comfortable boundaries but all the same requiring a pressure-free learning environment. As explained earlier within the WTC framework, learners are not truly forced to speak: rather, they are made to find convincing motivational factors to use the target language.

On the other hand several students noted that conversations can also be helpful among friends. One of them wrote: "I think speaking with my friends in French outside the classroom was most important because I could freely ask them questions about certain words without having to look

them up, or if neither of us knew how to express a certain idea, looking it up and both learning it together. This was most effective because we were not necessarily constrained by time limits or the structure of a classroom setting and could take our time, and we didn't judge each other for our mistakes or misuse of words like we might think students or teachers might do in a classroom". In this case, the learners displayed a high level of WTC without feeling obligated to speak. The difference may lie in different personalities. MacIntyre (2007) noted that extroverts thrive in unfamiliar situations whereas introverts do better in familiar settings. The first context could be equated to situations where one feels forced to speak (e.g., podcasting, recorded narrations, being immersed in a host country), and the second to more laid-back situations (low-stakes in class activities such as debates that were often not recorded, non-recorded article summaries, conversations with friends).

DISCUSSION

Grammar vs. Pronunciation

The between-group difference on what learners perceive to be correlates of high proficiency deserves attention. For the most part, learners of German believe that it takes a good control of grammar and vocabulary. Students of French, on the other hand, think that correct pronunciation is required to attain high oral proficiency. There are two possible explanations for this divergence. First, the difference may be attributed to teachers' influence. While neither course was geared specifically toward grammar or pronunciation, the German instructor is a certified OPI trainer and may have emphasized grammatical elements important to success on a German OPI more strongly. Mastery of verb tenses and connectors, along with repercussions on subject-verb order, is crucial at the Intermediate and Advanced levels targeted in that course. In comparison, French link words carry little consequence on syntax. The French instructor is also a phonetician, and her interest in pronunciation may have influenced students. In addition, native English speakers may find French pronunciation and German grammar/vocabulary more challenging than the other way around, as German phonology is closer to that of English. But German grammar, with distinct word order and complex declensions, may seem more daunting than the French system.

Recorded vs. not recorded

In both groups, activities that included recording technology were ranked as helpful to increase oral proficiency. But contrary to the French group the German group cited self-monitoring as a tool for improvement, explaining that listening to one's own speech induced noticing and subsequently correcting mistakes. The students of French appeared to favor conversation practice in a low-stress environment, i.e., ungraded debates that were not subject to self-evaluations. This finding supports the Affective Filter Hypothesis (Krashen, 1982): stressful situations raise the affective filter, preventing the processing of input; stress-free situations lower the filter, giving access to comprehensible input, which leads to learning. Yet students thought that being forced to tackle unfamiliar topics – an ability linked to higher proficiency on the OPI – was nonetheless positive. In summary, the best overall teaching strategy seems to involve a little pressure to provide a motivating challenge and increase the WTC in extrovert students, add low-stakes speaking practice to lower the affective filter and increase the WTC in introvert students, while giving opportunities for metacognitive development to facilitate noticing and accelerate second language acquisition.

Pedagogical implications

Preliminary results suggest that recording technologies paired with self-monitoring and self-evaluation tasks do benefit learners' metacognitive behaviors as respondents commented on the usefulness of noticing mistakes in order to correct them more easily. However, traditional conversations in and outside of class were found to be helpful too. Results varied across classes. Learners of French may generally ask for more pronunciation instruction but learners of German may request more grammar. Recording activities should perhaps emphasize different skills depending on the target language. Ducate & Lomicka (2009) showed that podcasting contributes to improving L2 pronunciation, so podcasting may be an appropriate option to develop that aspect in French courses. Podcasting could perhaps be utilized in German classes with more focus on grammar.

In addition to that, sequencing the skills differently for French and for German students with guidance for self-evaluation might reduce anxiety and satisfy the varying learners' needs. For instance, students could engage in three different recorded activities out of several practices for each: a narration, a debate, and a conversation. In the French group pronunciation could be the focus of the first, followed for the second by pronunciation and grammar, and finally pronunciation, grammar, and vocabulary for the third. The order in German class would be grammar, vocabulary, and lastly pronunciation.

There may also be anxiety linked to unfamiliar technology. The creation of podcasts might sound intimidating to some learners. In reality a podcast does not have to be anything more than a collection of speech samples. Breaking it down in those terms may relieve some of the learners' stress. MacIntyre (2007) remarked the complexity of the choices to evaluate before reaching a WTC level high enough to speak. Anxiety and personality types are major contributing factors, and the familiarity of the situation comes into play as well. He commented: "As expected, the extraverts showed higher WTC than the introverts when studying new L2 vocabulary words in a moderately unfamiliar situation. However, the pattern was reversed when the study conditions were highly familiar. Under that condition, the introverts actually showed higher WTC than the extraverts did" (p. 570). Providing a balance between familiar and unfamiliar learning situations seems crucial to cater to different personalities. With technology it may thus be important to schedule multiple sessions so that recording activities eventually become familiar, which will help introvert learners. A conversation course could be turned into an amateur French radio show. Learners would come to class to prepare the show on a regular basis throughout the semester, record a few episodes during the term (i.e., narrations, debates, conversations), and have the option to share the episodes with classmates, friends, or any target audience of their choice. Moreover, since participants in this study felt that immersion is the best way to achieve a higher proficiency level – and research supports their view (Lindseth, 2010) – stepping into the lab would equate stepping into a recording studio, transforming part of the course into a mini-immersion experience.

Having students reflect on their personalities so they can be made aware of how they learn and what situations will increase their WTC can also be a good time investment. Instruments are available to increase metacognitive awareness. There are Likert-scale questionnaires that can be distributed to students to guide their reflection. One of them is the Metacognitive Awareness Inventory (Schraw and Dennison, 1994). It contains 52 items, among which "I summarize what I've learned after I've finished", "I am a good judge of how well I understand something", or "I change strategies when I fail to understand". Those types of questions generate reflection on

personal behavior when facing challenges while describing strategies to cope with those problems.

CONCLUSION

As participants pointed out, some learners need to be pushed outside of their comfort zone and rise to the challenge of spontaneously speaking a second language. Others do better in more familiar environments. No matter the situation, in order to correct mistakes and improve L2 proficiency it is important for students to (1) be explicitly or implicitly convinced to speak, i.e., reach a high WTC, (2) become conscious of errors, i.e., develop noticing skills, (3) be aware of learning strategies, i.e., increase metacognitive awareness. Recording technology and traditional teaching methods combined together are able to provide opportunities for all three aspects and lead to increased oral proficiency.

ACKNOWLEDGEMENTS

ABOUT THE AUTHOR

Jessica Sertling Miller (Ph.D., University of Illinois at Urbana-Champaign), a native of Strasbourg, France, is an Associate Professor of French at the University of Wisconsin-Eau Claire. She teaches all levels of French courses, from beginner to linguistics and advanced conversation. She also supervises students in the French Education program. She has most recently published work on foreign language pedagogy and co-authored a textbook on French pronunciation, *Sons et Sens* (2013).

Jessica S. Miller
University of Wisconsin-Eau Claire
Foreign Languages, Hibbard 378
Eau Claire, WI 54702-4004
Home: (715) 254-1347
Office: (715) 836-4267
E-mail: millerjs@uwec.edu

REFERENCES

ACTFL/NCATE program standards for the preparation of foreign language teachers. (2002). Yonkers, NY: American Council on the Teaching of Foreign Languages.

Anderson, N. J. (2008). Metacognition and good language learners. In C. Griffiths (Ed.) *Lessons from good language learners* (99-109). Cambridge University Press.

Chinnery, G. M. (2006). Going to the MALL: Mobile assisted language learning. *Language Learning & Technology, 10*(1), 9-16.

Cooper, M. M., and Sandi-Urena, S. (2009). Design and validation of an instrument to assess metacognitive skillfulness in chemistry problem solving. *Journal of Chemical Education 86* (2), 240-245.

Ducate, L. & Lomicka, L. (2009). Podcasting: An effective tool for honing language students' pronunciation? *Language Learning & Technology, 13*(3), 66-86.

Graham, S. (2006). A study of students' metacognitive beliefs about foreign language study and their Impact on learning. *Foreign Language Annals, 39*(2), 296-309.

Henn, S. (2012). On Utah's 'Silicon Slopes,' Tech Jobs Get a Lift. *NPR* Retrieved from www.npr.org/2012/03/12/148252561/on-utahs-silicon-slopes-tech-jobs-get-a-lift

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Lord, G. (2008). Podcasting communities and second language pronunciation. *Foreign Language Annals, 41*(2), 364-379.

Lindseth, M. U. (2010). The development of oral proficiency during a semester in Germany. *Foreign Language Annals, 43*(2), 246–268.

McCarty, S. (2005). Spoken Internet to go: Popularization through podcasting. *The JALT CALL Journal, 1*(2), 67-74.

McCroskey, J. C., & Richmond, V. P. (1991). Willingness to communicate: a cognitive view. In M. Booth-Butterfield (Ed.). *Communication, cognition, and anxiety*, 19-37.

MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The Modern Language Journal*, 91, 564–576.

Miller, J. S. & Hlas, A. (2010). The possibilities and potential of podcasting. In M. Bloom and C. Gascoigne (Eds.). *Central States Conference Report 2010*, 27-44.

O'Bryan, A., & Hegelheimer, V. (2007). Integrating CALL into the classroom: The role of podcasting in an ESL listening strategies course. *ReCALL, 19*(2), 162-180.

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

Onsrud, E. M. (2009). *Podcasting in the foreign language classroom: Improving communication.* (Unpublished Master's thesis). University of Wisconsin-Stout, Menomonie, Wisconsin.

Rivers, W. P. (2001). Autonomy at all costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners. *The Modern Language Journal, 85*(2), 279-290.

Rosell-Aguilar, F. (2007). Top of the pods: In search of a podcasting "podagogy" for language learning. *Computer Assisted Language Learning, 20*(5), 471-492.

Schraw, G. and Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19,* 460-475.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26,* 113-125.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition*, 235-256. New York: Newbury House.

Tanner, K. D. (2012). Promoting Student Metacognition. *Cell Biology Education—Life Sciences Education, 11* (Summer): 113–120.

Thompson, D. R. (2012). Promoting metacognitive skills in intermediate Spanish: Report of a classroom research project. *Foreign Language Annals, 45*(3), 447-462.

Thorne, S. L., & Payne, J. S. (2005). Evolutionary trajectories, Internet-mediated expression, and language education. *CALICO Journal, 22*(3), 371-397.

ENGLISH L2 VOWEL ACQUISITION OVER SEVEN YEARS

Murray J. Munro, Simon Fraser University

Tracey M. Derwing, University of Alberta

Kazuya Saito, Waseda University

Although cross-sectional research designs have been widely used in the evaluation of L2 phonetic learning, longitudinal studies of L2 speech production are rare. As a result, it is difficult to draw strong conclusions about the effects of language experience on L2 phonetic acquisition. This investigation of adult Slavic (Russian and Ukrainian) and Mandarin speakers tracks their English high vowel productions during seven years of residence in an English-speaking area. At the outset of the study, all participants had limited English oral proficiency. To evaluate phonetic learning, recordings of English vowels produced in controlled phonetic contexts were compared at the outset of the study, at one year, and at seven years. Vowel intelligibility was assessed through listener judgments in a blind identification task, and vowel accuracy was evaluated through acoustic measurements. While the results support the proposal that adults remain open to phonetic learning, they also indicate a dramatic slowing of the acquisition process by the end of the first year.

INTRODUCTION

Despite decades of research on adult second language (L2) phonetics, many aspects of the temporal development of L2 vowel and consonant production remain poorly understood. On the one hand, researchers have established that adults commonly do not learn to produce fully native-like L2 segments even after years of exposure to the L2. On the other hand, researchers have gained only limited insights into the amount of learning that actually does occur as a function of L2 experience and the time course of that learning (as opposed to instructed pronunciation learning). As part of an extensive longitudinal project examining oral language acquisition (see Derwing & Munro, 2013; Derwing, Munro, & Thomson, 2008), the present study aims at uncovering new details about the vowel acquisition process in adult ESL learners.

A much-needed type of work within the field of L2 phonetic learning is longitudinal research. Despite the preponderance of cross-sectional studies of segmental production comparing learners with different lengths of L2 residence (LOR), few solid conclusions about learning trajectories can be drawn, partly because LOR is a poor measure of L2 experience and also because of the problem of confounding factors in samples of learners drawn from different LOR populations. A further issue is that much of the work on segmental production has focused on phonetic "accuracy" – the extent to which L2 segments match those of the target L1 speakers. While that orientation can be useful in testing certain theoretical models, it tends to be misleading, and even counter-productive in applied phonetics, because phonetic accuracy is not a prerequisite for speech intelligibility (Derwing & Munro, 2009). For instance, inaccurate production of certain consonants such as /θ/, which participates mainly in low functional load phonemic distinctions, may have few serious communicative consequences for English

L2 speakers (Brown, 1991, Munro & Derwing 2006). Furthermore, if a segment such as /oʊ/ is produced intelligibly, but not entirely accurately, an improvement in accuracy may not yield any comprehension benefits for a speaker's interlocutors. In summary, a lack of acquisition of certain phonetic dimensions may not hamper L2 learners' communication skills, and improved accuracy on other dimensions may be communicatively irrelevant.

Although focused instruction is known to benefit L2 segmental production (e.g., Saito & Lyster, 2012), data on the time course of L2 phonetic development may be of use in identifying the types of learning that typically occur without intervention. Speech phenomena that fossilize early without instruction may be good candidates for classroom attention, provided they have a clear impact on intelligibility. At present, the results of studies of adult phonetic acquisition over time are mixed. Some research suggests that beyond a brief initial period of rapid phonetic learning, further L2 exposure has only limited effects on phonetic accuracy. Flege (1988), for instance, observed no cross-sectional difference in global foreign accent ratings between Taiwanese speakers of English with 1 year of US residence and those with 5 years of residence. Furthermore, Derwing and Munro (2013) found no longitudinal improvement in global accent ratings of Mandarin or Slavic speakers between their second and seventh year of Canadian residence. However, ratings of the Slavic speakers' comprehensibility (easy vs. difficult to understand) did improve significantly over the same interval. When specific segments have been considered, such as Japanese speakers' productions of English /ɹ/, learners with greater LOR have sometimes outperformed shorter-term residents (Flege, Takagi, & Mann, 1995) and sometimes not (Larson-Hall, 2006).

The current study aims to establish the learning trajectories for a particular set of segments – in this case high vowels – in English L2 learners after arrival in an English-speaking country. Here we address *intelligibility* through listener identifications and *accuracy* through acoustic measurements. This investigation extends earlier longitudinal work (Munro & Derwing, 2008) examining the vowel development of Mandarin and Slavic (Russian and Ukrainian) speakers during their first year in Canada. On arrival, all speakers in that study had low oral proficiency, and all were students in the same ESL program – one featuring no focused pronunciation instruction. Their productions of ten different vowels in CVC contexts were evaluated by both phonetically trained and untrained listeners. Significantly improved vowel intelligibility was observed during the first 6–8 months of Canadian residence, followed by a leveling off. This outcome appears to conform to Flege's (1988) proposal for global foreign accent. However, the same study revealed continued improvement over the entire year in both groups' performance on /ɪ/, which is missing from the phonemic inventories of both groups' L1s. Furthermore, /ɪ/ proved to be the least intelligible L2 vowel at both the beginning (5–31% correct) and the end (21–48%) of the study, with both groups of speakers tending to pronounce it as /ɛ/. Intriguingly, /ʊ/, which also does not occur in the speakers' L1s, showed somewhat better intelligibility than /ɪ/ at the outset, but no evidence of improvement over the year. A further finding was that vowels were produced more intelligibly in bVC than in pVC contexts, perhaps because of greater word frequency for the bVC items, which might lead to more exposure to native exemplars and, in turn, to more accurate perceptual representations on the part of the learners. Some questions that remained unanswered were whether further improvement would occur on /ɪ/ after the end of the first year and whether performance on /ʊ/ would remain unchanged. Another issue was whether the discrepancy in performance on bVC vs pVC would persist beyond the first year.

In view of the above outcomes from Munro and Derwing (2008), we focus our attention on the following research questions:

> Q1. In the absence of pronunciation instruction, to what degree will the adult learners of English improve in high vowel intelligibility and accuracy between years 1 and 7 of residence in their L2 environment?

Q2. If improvement occurs, which specific vowels will be affected?

Q3. Will the difference in intelligibility favoring bVC over pVC persist after 7 years of residence?

METHODS

Speakers

The speakers were 13 Mandarin and 18 Slavic-speaking adults who participated in Munro and Derwing (2008), all of whom were recent arrivals in Canada with low oral proficiency at the outset of the study. All had enrolled in the same ESL program, which featured no focused pronunciation instruction. Further details are given in Derwing, Munro & Thomson (2008) and Derwing and Munro (2013).

Speech Materials

Digital recordings were made of 10 bVC and 10 pVC productions from each speaker, where V = / i ɪ eɪ ɛ æ u ʊ oʊ ɑ ʌ /, and C = /t/ (except /k/ in the case of /bʊk/). At a number of testing points over the course of the study, tokens were elicited via a delayed repetition task in which the speakers heard the target words produced in the frame "The next word is ___," and responded with "Now I say ___." From the original recordings, the target words were excised from the sentence frame, normalized for peak amplitude, and saved as individual audio files for randomized intelligibility assessment and acoustic analysis. Although the speaking task was completed at multiple testing points, for the purposes of the present study, we will focus on productions from the one-year and seven-year points, and will compare them with productions from the outset of the study. Furthermore, we will be concerned only with the English high vowels / i ɪ u ʊ /.

Intelligibility Assessment

Vowel intelligibility was assessed by three phonetically-trained native English listeners, two of whom were the first two authors. The listeners heard the productions through headphones during a blind, randomized ID task in which they identified the closest native English vowel to the one heard. Items were presented via computer playback software (22.05 kHz, 16 bits), and identifications were made with screen buttons labeled with phonetic symbols. Buttons for all the possible vowel targets were available, along with 'replay' and 'unknown' buttons. Because of the large number of tokens, judgments were completed over several sessions.

Acoustic Assessment

For the high front and high back vowels, fundamental frequency (F0) and first and second formant (F1 and F2) measurements in Hz were made at the vowel midpoints from pitch tracks and formant tracks obtained through linear predictive coding in Praat (Boersma & Weenink, 2012). These were converted to Bark values, which correspond more closely than Hz measurements to human perception. Bark difference values (F1-F0, F2-F1) were then computed to reduce between-speaker differences, including gender effects, resulting from variability in vocal tract size.

RESULTS

Intelligibility

Identification data were converted to %-correct identification scores by tallying the number of times the high vowel tokens were labeled as the target vowel. Figure 1 provides mean identification scores pooled over /i ɪ u ʊ/ for the Mandarin and Slavic groups at two test times, with results broken down by initial consonant. These data were submitted to a mixed-design Analysis of Variance with first language (L1) as a between-groups factor and Time (T = 1 year, 7 years) and Initial Consonant (IC = b,

p) as within-group factors. Only the effect of IC proved statistically significant, $F(1, 29) = 4.685$, $p = .039$, $\eta_p^2 = .139$, indicating that high vowels in words beginning with /b/ were produced more intelligibly than those in words beginning with /p/. All other effects and interactions missed significance at p > .1
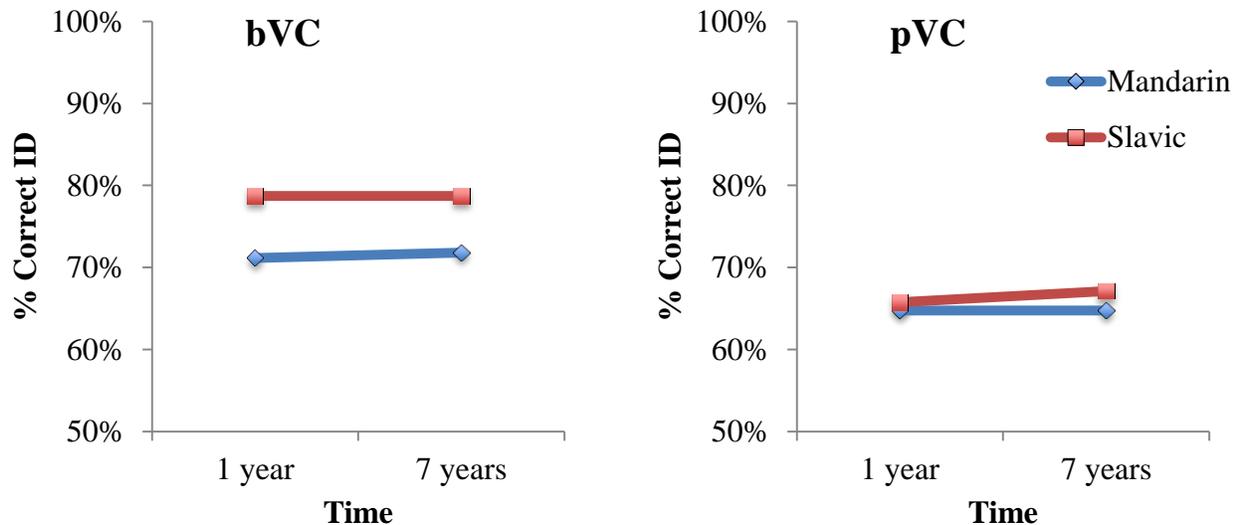
Figure 1: Mean intelligibility (% correct ID) of the two groups' high vowel productions at one year and seven years for bVC and pVC words.

Although the ANOVA provided no indication of an *overall* improvement in high vowel intelligibility, it is still possible that performance improved significantly on one or more vowels, but not on the others. Ideally, we would have liked to carry out statistical analyses for each vowel, but small cell sizes made such an approach inappropriate. Therefore, we present here an informal comparison across vowels in bVC context only, which is illustrated in Figure 2. To provide a fuller context, we include intelligibility data from the outset of the study (Munro and Derwing, 2008) with scores from the one-year and seven-year points.

For /ɪ/, improved intelligibility between the outset and the one-year point appears to have occurred in both groups, but there is no indication of meaningful improvement after one year. However, for /ʊ/, both speaker groups appear to show higher intelligibility at seven years. There is also slight improvement by the Slavic speakers on /i/, and by both groups on /u/.
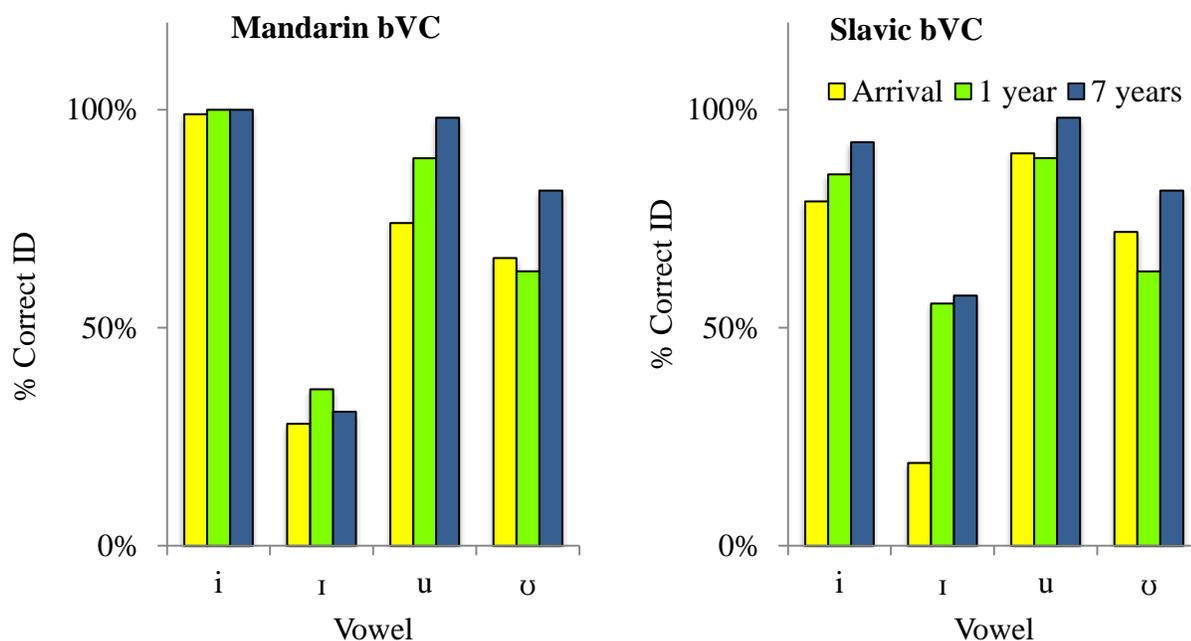
Figure 2: Mean intelligibility (%-correct identifications) of high vowels in bVC context produced by the Mandarin (left) and Slavic (right) groups at three times.

Acoustic Properties

To probe further the acquisition of the four vowels, acoustic data were informally evaluated. Mean values for the two groups are presented in Figure 3, with F2–F1 (Bark) on the x-axis and F1–F0 (Bark) on the y-axis. Bark scaling allows us to interpret the figure as an approximate representation of vowel height and advancement, with the arrows representing the direction and extent of change in tongue position for each vowel from the beginning of the study (in yellow) to the end of one year (in green) and until year 7 (in blue). Data were pooled from the bVC and pVC productions and are therefore not fully comparable with Figure 2. For both speaker groups, the clearest indication of change was for /ɪ/, which became higher and more advanced, particularly over the first year. Since that vowel was typically misproduced as /ɛ/, the direction of change is the expected pattern for improved intelligibility. The Slavic group, and to a much lesser degree the Mandarin group, showed some additional change in the production of /ɪ/ between the first and seventh years. Change in /i/ was also considerably greater in the Slavic than the Mandarin group, a finding that fits well with the near-ceiling intelligibility on that vowel by the Mandarin speakers, and improved intelligibility over time by the Slavic speakers. For /u/, the Mandarin speakers showed higher and more forward articulations during the first year, followed by a regression back to the original vowel position, while the Slavic speakers showed higher and more forward productions during year one, followed by no further change. Finally, for /ʊ/, the Mandarin group showed lower, more back articulations over the first year, followed by a regression, while the Slavic group exhibited higher, more forward productions during year one, followed by a slight movement to higher positions.
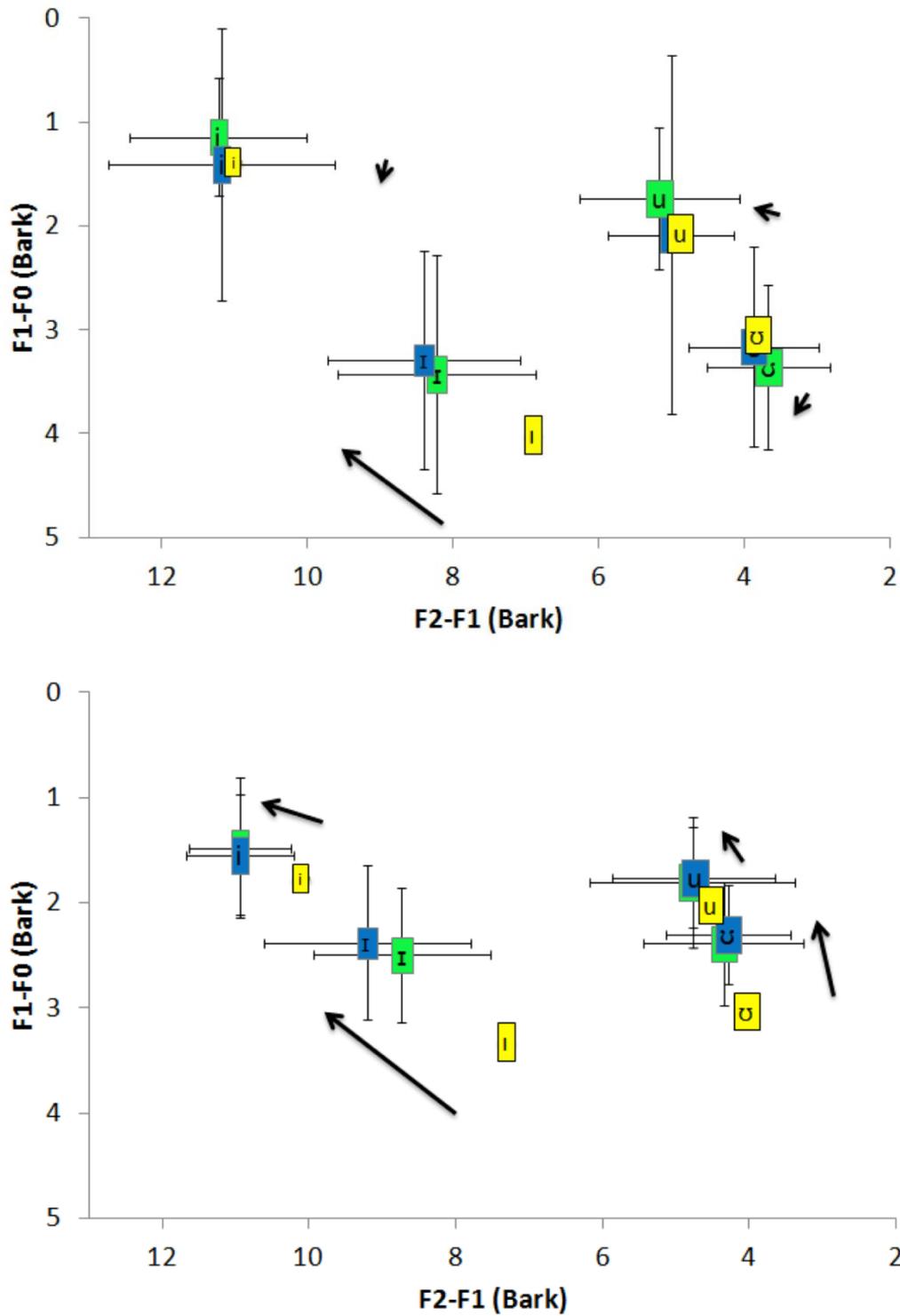
Figure 3: Changes in vowel formant frequencies for the Mandarin (top) and Slavic (bottom) groups from the outset of the study (yellow), to the 1-year point (green) and the 7-year point (blue).

DISCUSSION

In response to our original research questions, we can now make several observations.

Q1. Improvement in high vowel intelligibility and accuracy between years one and seven

Although there was no statistically significant change in overall high vowel intelligibility, an examination of the individual vowels suggests that some improvement in intelligibility did occur in both groups of learners. The acoustic data support this interpretation to some degree in that some productions appear to have improved in accuracy.

Q2. Specific vowels showing improvement between years one and seven

Listener evaluations suggest that the intelligibility of the Slavic speakers' /ɪ/ productions improved between years one and seven, while the acoustic data indicate improved accuracy (higher, more advanced productions) over the same time period. Data from the Mandarin speakers on /ɪ/ were not consistent with any improvement. While both groups appeared to show more intelligible productions of /u/ and /ʊ/ after seven years, the acoustic data suggested a regression in the Mandarin group toward original values. It is difficult to interpret these findings, partly because measurements of vowel formants taken at a single point do not reflect all the aspects of a vowel that determine its quality. (Vowel-inherent formant movement, for instance, is not captured.) A more detailed examination of the acoustic data in connection with specific vowel exemplars may lead to further understanding of the changes that the speakers actually implemented in their production strategies.

Q3. Intelligibility advantage for bVC vowels over pVC vowels

As was the case at the end of year one, vowels in the pVC context continued to be less intelligible than bVC vowels at the seven-year point. Whether this is a permanent aspect of L2 speech production should be explored through an examination of the productions of longer-term residents. A detailed study of a large set of vocabulary items varying in frequency should be conducted to determine whether there is a lexical effect for phonetic acquisition.

CONCLUSION

This longitudinal investigation of the acquisition of English vowels provides support for the view that the largest gains in segmental intelligibility and accuracy occur during the first year of residence in an L2-speaking area. However, the findings also suggest that phonetic learning does not cease altogether at the end of the first year. Rather, further improvements in intelligibility and refinements to production accuracy may occur naturalistically in some segments during the years that follow.

ACKNOWLEDGMENTS

ABOUT THE AUTHORS

A former ESL teacher, Murray Munro is a Professor of Linguistics at Simon Fraser University. His applied phonetics research has appeared in a variety of international journals. Email: mjmunro@sfu.ca

Tracey Derwing is a Professor of TESL in the Department of Educational Psychology at the University of Alberta. She has conducted numerous studies of pronunciation and oral fluency development in L2 learners. Email: tracey.derwing@ualberta.ca

Kazuya Saito is an Assistant Professor of English in the School of Commerce at Waseda University (Tokyo, Japan). His research focuses on the role of experience and age in L2 speech learning in both naturalistic and instructed settings. Email: kazuya.saito@waseda.jp

Correspondence concerning this article should be addressed to Murray J. Munro, Department of Linguistics, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6, mjmunro@sfu.ca.

REFERENCES

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program]. Retrieved July 2012 from http://www.praat.org/

Brown, A. (1991). Functional load and the teaching of pronunciation. In Brown, A. (Ed.), *Teaching English Pronunciation: A Book of Readings* (pp. 221–224). Routledge, London.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching 42*, 276–490.

Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A seven-year study. *Language Learning*, *63*, 163-185.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics 29*, 359–380.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America 84,* 70-79.

Flege, J. E., Takagi, N., & Mann, V. A. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. *Language and Speech 38*, 25–55.

Larson-Hall, J. (2006). What does more time buy you? Another look at the effects of long-term residence on production accuracy of English /ɹ/ and /l/ by Japanese speakers. *Language and Speech 49*, 521–548.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning 58*, 479–502.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System 34*, 520−531.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning 62*, 595–633.

THE DEVELOPMENT OF /θ/, A VARIABLE GEOGRAPHIC PHONETIC FEATURE, DURING A SEMESTER ABROAD: THE ROLE OF EXPLICIT INSTRUCTION

Angela George, Kennesaw State University

Previous studies have shown positive effects of instruction on pronunciation after a time abroad (Lord, 2010). Four studies have examined the production of /θ/ in Spanish, (Geeslin & Gudmestad, 2008; Henriksen, Geeslin, & Willis 2009; Knouse 2013, and Ringer-Hilfinger, 2012) and one found that eight out of nine learners increased their production during the 7-week immersion in León, Spain in the absence of explicit instruction (Henriksen, Geeslin, & Willis 2009). The current study differs from previous work by examining the effect of explicit instruction on the development of /θ/, a salient dialectal feature to North-Central Spanish, in university-level adult learners of Spanish during a semester abroad in North-Central Spain. After completing a pretest, the experimental group, consisting of 14 students, received explicit instruction on when North-Central Spaniards use /θ/, while the control group, consisting of 10 students, did not. Surprisingly, the experimental group decreased their /θ/ use on all three tasks from the beginning to the end of the semester, while the control group increased their /θ/ use over time on all three tasks. Instruction alone is not sufficient, and other factors such as attitude toward this dialect, exposure to the dialect, and proficiency level play a role.

INTRODUCTION

The current study addresses the effect of instruction on the acquisition of the interdental fricative /θ/, a geographically variable Castilian[1] Spanish dialectal phonological feature. It examines two groups of students studying abroad for one semester in Northern-Central Spain, one that received explicit formal instruction on when to use the feature, a defining feature of the dialect (Hualde, 2005), and one that did not. In Castilian Spanish, /θ/ is used for orthographic 'z' (e.g., **z**apato) and 'c' before 'i' and 'e' (e.g., gra**c**ias, ha**c**er), whereas in Latin American Spanish, only /s/ is utilized in these contexts. In English, /θ/ occurs with different orthographic symbols, but similar to Castilian Spanish, can occur word initially, medially, or finally.

BACKGROUND

Phonological Acquisition and Spanish Phonetics Instruction

A handful of studies demonstrate that taking a university-level Spanish phonetics course leads to improved pronunciation of specific non-dialectal features. Castino (1996) and Lord (2005) found improvement of consonant allophones in students taking a Spanish phonetics course. Learners of Spanish exhibited increased pronunciation ratings as judged by native speakers of Spanish after participating in a podcast project to improve

---

[1] Castilian Spanish refers to the Spanish spoken in Northern-Central Spain (Hualde, 2005).

their pronunciation (Lord, 2008). Lord (2010) found that Voice Onset Time of /b, d, g/ improved for four learners of Spanish who received Spanish phonetic instruction prior to studying abroad in Mexico.

Study Abroad and the Acquisition of Dialectal Features

Several studies that investigated the acquisition of phonological dialectal features have shown movement toward more target-like norms. In an investigation on the effect of instruction on the perception of Andalusian[2] Spanish, the experimental group improved significantly on the comprehension of sentences containing synalepha, the linking or elision of vowels across word boundaries, after receiving explicit phonetic instruction on Andalusian dialectal features (Rasmussen & Zampini, 2010). Production studies have also found movement toward more target-like dialect features (e.g., Henrikson, Geeslin, & Willis, 2010). Willis, Geeslin, and Henriksen (2009) found more target-like productions of /θ/ after a 7-week study abroad program in eight out of nine high school learners of Spanish. However, Ringer-Hilfinger (2012) found only two learners produced [θ] a total of six times (out of 209 possible contexts). One learner produced [θ] six months after returning from studying in Madrid due to a strong desire to sound like a speaker of Castilian Spanish. The other learner produced [θ] prior to studying abroad, possibly because she had a Castilian Spanish speaking instructor in high school with whom she traveled to Spain for one week. Geeslin & Gudmestad (2008) found that nine out of 130 learners living in the U.S. produced [θ]. They concluded that contact with native speakers upon returning from study abroad was more important than the study abroad experience. Knouse (2013) compared an At Home (AH) group taking Introduction to Hispanic Linguistics with a native Castilian Spanish-speaking instructor to a Study Abroad (SA) group in Salamanca and found that, at the end of 6 weeks, only six students in the SA group produced [θ] a total of 33 times in a reading passage and semi-spontaneous speech[3]. The Goldvarb analysis, which predicts which factors had a statistically significant effect on [θ] production, showed that [θ] was favored with the grapheme 'z' over 'c'. It was also favored by intermediate and beginner learners, and disfavored by advanced learners, contrary to Geeslin & Gudmestad (2008). Spanish majors and minors favored the use of [θ] over other students.

Variation in Second Language Acquisition
The variationist model for second language acquisition (SLA) based on Preston's (2000, 2002) and Fasold and Preston's (2006) sociolinguistic model explains interlanguage (IL) variation in adults who started learning the TL as adults or older adolescents. According to this model, IL variation is caused by sociocultural factors, linguistic context, and time; forms learned earlier are more internalized and therefore more automatic, while forms learned later require more attention and control (Preston, 1989).

The present study seeks to answer the following research questions: What is the effect of explicit instruction while studying abroad in North-Central Spain on the L2 production of the voiceless interdental fricative in Spanish? What individual or social factors explain this effect?

---

[2] Andalusía is a region in southern Spain.
[3] The semi-spontaneous speech consisted of responding out loud to questions written in Spanish.

DATA COLLECTION AND ANALYSIS

Participants

Ten students in the experimental group (EG) received explicit instruction on when to use [θ] as part of the course, Introduction to Hispanic Linguistics, taught by a local Castilian-Spanish professor. Fourteen students in the control group (CG) did not take this course, and did not receive explicit instruction on when to use [θ].

All participants were native speakers of English studying abroad for one 13-week semester in Toledo, Spain. They mostly ranged in age from 18-21 (one participant was 31). Six students were male and 18 were females. All attended various universities and colleges in the U.S., majoring or minoring in Spanish, and all had completed intermediate Spanish prior to studying abroad. In the EG, 1 student was in her second year, 5 in their third, and 4 in their fourth or fifth year. In the CG, 11 students were in their third year and 3 were in their fourth year. The average length of formal study of Spanish was 7.40 years (range 5-10 years) for the CG and 6.30 years (range 4-9) for the EG. This difference was not statistically significant ($t(23) = 1.184$, $p = .249$), implying the two groups were similar in terms of prior Spanish instruction. The participants were told they were part of a study investigating language development as a result of study abroad.

Instruction

In the third week of the semester abroad, during the Introduction to Hispanic Linguistics course taught twice a week for 1.25 hours, students were taught the contexts in which /θ/ is produced in Castilian Spanish spoken in North Central Spain. They contrasted this with other varieties of Spanish that do not use /θ/. Students were informed that a written 'z' and 'c' before 'e' or 'i' in Castilian Spanish is pronounced as [θ], whereas in Latin American dialects it is pronounced [s]. They completed in-class activities for 30 minutes, consisting of identifying the contexts where /θ/ could be produced in Castilian Spanish, practicing the articulation of /θ/, and listening to songs to identify where /θ/ was used, if it was used. They were given the opportunity to practice producing /θ/ with a partner after identifying the contexts in a text where /θ/ could be produced. Homework was assigned to practice identifying the contexts where /θ/ is used in Castilian Spanish. Halfway through the semester, they were tested on this as part of an in-class open-book written midterm exam. Students in the EG were told the contexts where they could use [θ], but were not necessarily told that they should use it in their daily speech.

Assessment

The pretest and posttest consisted of students completing the same three tasks, once during the second week and once during the twelfth week of the semester. The first task, a conversation with a native speaker of Castilian Spanish, was designed to elicit spontaneous speech. The participants might have bee more likely to use the /θ/ if they were speaking with a native speaker who also exhibited the feature. The final two tasks, a reading passage and a word list, were designed to elicit more carefully monitored speech. The reading passage contained the words from the word list and was performed first. In the word list, participants read one index card at a time. There were 60 index cards and 23 possible contexts where [θ] could be used. The participants completed these two tasks

individually in a quiet room and were told to pretend as if the native Castilian Spanish speaker was in the room.

Instruments

At the beginning and end of the semester, the students completed a Language Contact Profile based on Freed, Dewey, Segalowitz, and Halter (2004), where they reported how many days per week and how many hours per day they interacted with native speakers of Spanish and engaged in noninteractive activities such as watching television. The participants also completed a questionnaire, based on Drummond (2010), designed to elicit attitude toward Castilian Spanish, motivation to learn Spanish, awareness of the Castilian Spanish dialect, desire to sound Spanish, and anxiety toward Spanish pronunciation. Finally, students completed information based on Qui (2011) about their Spanish-speaking social networks, which included information on the dialects of the speakers in their network and the type and amount of interaction with these speakers.

Participants completed a semi-structured interview with the researcher at the beginning and end of the semester, in the language of their choice, in order to obtain information about previous travel to Spain, previous instructors of Spanish, travel throughout the semester, and activities with Spanish speakers throughout the semester. This data was not used to calculate how often the participants used [θ].

Procedure

Participants were recorded conversing with a native speaker of Spanish, for an average of eight minutes, in a small, quiet room. Each participant then went to another room and read a passage followed by a word list. The researcher started the digital recorder and then left each participant alone in the room. Four native speakers, not included in the analysis, also read the reading passage and word list, producing [θ] in all possible contexts.

The researcher transcribed the recordings, specifically listening for [θ]. Three non native and one native speaker of Spanish, all advanced graduate students in Hispanic Linguistics with training in Hispanic phonology and phonetics, listened to a sample of 174 sounds where [θ] could have been produced. Variants identified by each rater were [θ], [s], and [z]. A reliability test showed a high degree of inter-rater reliability[4].

RESULTS

The total number of productions of [θ] was divided by the number of possible contexts to calculate a percentage of [θ] use for each individual. Table 1 shows the percentages of [θ] used in each task as well as all tasks combined before and after the instruction period for the EG, and Table 2 shows the same information for the CG.

---

[4] Cronbach's alpha was 0.89, which is greater than 0.70, offering evidence of inter-rater reliability.

Table 1
Experimental Group's (N=10) [θ] production

| Task | Before | | | After | | |
	Tokens (#)	Mean (%)	Standard Deviation (%)	Tokens (#)	Mean (%)	Standard Deviation (%)
Word List	15/215	6.98	19.03	9/229	3.93	6.30
Reading Passage	3/353	0.85	1.89	3/361	0.83	2.64
Spontaneous Speech	4/214	1.87	2.97	2/253	0.79	1.73
All tasks combined	22/782	2.81	6.33	14/843	1.66	2.90

Table 2
Control Group's (N=14) [θ] production

| Task | Before | | | After | | |
	Tokens (#)	Mean (%)	Standard Deviation	Tokens (#)	Mean (%)	Standard Deviation
Word List	37/171	21.64	32.39	66/305	21.57	34.80
Reading Passage	33/503	6.56	15.96	48/504	9.52	21.20
Spontaneous Speech	4/308	1.30	3.50	16/325	4.92	7.31
All tasks combined	100/982	10.18	16.01	130/1134	11.46	18.84

With the exception of spontaneous speech at the beginning of the semester, the CG produced [θ] more than the EG in every task before and after the instruction. Despite this trend, the results of the ANOVA indicate that the CG group produced [θ] significantly more than the EG only in spontaneous speech at the end of the semester $F(1, 23) = 8.33$, $p = .008$.

Several variables may account for the fact that the CG produced [θ] more than the EG. The amount of self-reported contact in Spanish at the end of the semester was on average 365 contact hours for the CG and 212 contact hours for the EG. This difference was significant, $t(23) = 2.186$, $p = .039$; however, after the Bonferroni correction, it is not significant. The other variables show trends that favor [θ] production in the CG, but none were statistically significant. The Castilian Spanish social network strength; Spanish proficiency, as self-reported at the beginning of the semester; the number of years of

formal Spanish education; and the number of weekend trips spent in the host country were greater in the CG than the EG. The EG travelled more with Puerto Rican students on the weekends than the CG. All of these factors could account for why the CG produced more [θ] than the EG.

Seven variables were measured based on a six-point Likert scale as suggested by Dörnyei (2003). Attitude toward Castilian Spanish, anxiety toward speaking Spanish, Castilian dialectal awareness, motivation to learn Spanish, and amount of Spanish learned were not significantly different between the EG and CG. Despite this, the attitude toward Castilian Spanish increased from the beginning to the end of the semester for the CG and decreased for the EG. At the end of the semester, the CG's desire to sound Castilian was significantly higher than the EG. This desire would presumably include producing [θ], a feature which students in both groups noticed.

The students in the EG produced fewer instances of [θ], which could be because of more contact with Puerto Rican Spanish during weekend travel. More students in the EG lived in the dorm where the L1 was commonly heard; thus they could have been exposed to less Spanish. In the CG, about half lived in the dorm and half lived with a family. At the beginning of the semester, about half of the students in each group preferred the sounds of Castilian Spanish, but at the end of the semester only 40% in the EG preferred Castilian Spanish versus 64% in the CG. Finally, 2 (out of 10) students in the EG and 12 (out of 14) students in the CG were previously exposed to Castilian Spanish either via a previous instructor or previous trip to Spain. These factors could contribute to the CG producing more instances of [θ].

One student in the EG mentioned the reason for studying in Spain was to learn Castilian Spanish as opposed to Latin American Spanish. All others stated that they wanted to travel in Europe. Similarly in Ringer-Hilfinger (2012), only one out of four students who had studied abroad and produced [θ] mentioned using [θ] in order to fit in with the local Castilian dialect.

Linguistic Variables

A one-way ANOVA was performed to determine the effect of linguistic variables (lexical frequency, grapheme, and word placement) on instruction. Lexical frequency was determined based on the *Corpus de referencia del español actual* (CREA) by searching oral documents in Spain between 1970 and 2012. Lexical frequency was calculated by totaling the number of high frequency words, or those containing 200 or more tokens in the corpus and then the number of low frequency words, or those containing fewer than 200 tokens. Instruction did not have an effect on lexical frequency as both the CG and the EG produced [θ] in similar amounts of low frequency words. The EG did not produce [θ] in any high frequency words.

It is possible for [θ] to be produced with the graphemes 'z', 'ci', or 'ce'. The CG produced [θ] significantly more than the EG with the grapheme 'ci' in the word list at the end of the semester, $F(1, 3) = 28.80$, $p = .013$. Similarly, the CG produced [θ] significantly more with the grapheme 'z' in the word list at the end of the semester, $F(1, 3) = 19.97$, $p = .011$. This aligns with Knouse (2012), who also found that six students studying abroad in Spain for 6 weeks favored the use of [θ] with the grapheme 'z' over 'c'.

Instruction did not have an effect on word placement, or whether [θ] was produced word initially or word medially. Both groups produced [θ] in word medial position about the same amount. The EG did not produce [θ] word initially.

DISCUSSION

Overall, the participants in both groups exhibited low uses of [θ]. This could be because, according to Preston's (2000) model, [s] was learned first as the vernacular form and [θ] was learned later as the post-vernacular form. According to Tarone (2002), despite hearing [θ] frequently in the input, the learners adopt voices that are socially important to them. Thus, when choosing between two forms, they choose which form is more important to them for a variety of social reasons. Geeslin and Gudmestad (2008) found almost categorical use of [θ] in near-native speakers of Spanish, but the learners in this study were most likely not advanced enough to produce [θ] categorically. In contrast, Knouse (2013) found that intermediate and beginning students of Spanish favored this feature over advanced students; however, no students produced [θ] categorically.

The students in the EG slightly decreased their usage over time, with half stating in the interview that they were not trying to sound like Castilian Spanish speakers, but were trying to use general correct pronunciation. Two students indicated that it was cognitively difficult to produce [θ] in spontaneous speech, while two preferred other varieties over Castilian Spanish. Finally, two students were not sure if they would use [θ] in the future, when they returned to the US, where Castilian Spanish is a minority variety.

The students in the CG increased their production of [θ] throughout the semester, with a significant increase in spontaneous speech. Thus, factors other than instruction help account for this increase. More specifically, the CG experienced significantly more Spanish contact, a significantly stronger desire to sound like a Castilian Spanish speaker, more previous exposure to Castilian Spanish, and fewer weekend trips with Puerto Rican students.

LIMITATIONS AND FUTURE DIRECTIONS

Several limitations are evident in this study. In addition to small sample size, the only proficiency measure besides the years of Spanish studied was a self-reported measure at the beginning of the semester. Also, the groups were not randomly assigned as students chose whether or not to take the course in which they received instruction. Future studies would need to make sure the groups are more equivalent as higher proficiency levels can lead to more contact with native speakers (Segalowitz & Freed, 2004). Future studies might assess whether all participants are aware of the rule for when to use /θ/ as well as investigate whether instruction has an effect on perception. Due to the fact that the instruction students receive in the US prior to studying abroad may influence their (non-)use of this feature, it would be useful to compare two groups of beginning students studying abroad, one that receives instruction and one that does not.

In the current study, instruction does not lead to increased production of [θ], but rather leads to decreased production of [θ]. The individual factors accounting for less use of [θ] by the EG are less contact in Spanish, less previous exposure to this dialect prior to studying abroad, a weaker desire to speak Castilian Spanish, and more weekend trips with Puerto Rican students.

ABOUT THE AUTHOR

Angela George earned her Ph.D. in Hispanic Linguistics from the University of Minnesota in May, 2013. Her doctoral thesis investigated the development of dialectal features in learners of Spanish during a semester abroad in Spain. Her research interests include increasing pragmatic and sociolinguistic competence in learners of Spanish in various contexts, such as the classroom, immersion (i.e., study abroad), and home (i.e., heritage speakers). She has taught a range of Spanish courses, from beginning and intermediate levels to advanced composition and communication courses and Introduction to Hispanic Linguistics. She has also taught foreign language education teaching methods courses. She is currently an Assistant Professor of Spanish and Foreign Language Education. Contact Information: Kennesaw State University, Department of Foreign Languages 1000 Chastain Road, MD 1804, Kennesaw, GA 30144. Email: ageorg30@kennesaw.edu

REFERENCES

Castino, J. (1996) Impact of a phonetics course on FL learners' acquisition of Spanish phonology, *Selecta: Journal of the Pacific Northwest Council on Foreign Languages*, *17*, 55-58.

Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in Second Language Acquisition*, *26*, 249-73.

Dörnyei, Z. (2003) *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Drummond, R. (2010). Sociolinguistic variation in a second language: The influence of local accent on the pronunciation of non-native English speakers living in Manchester. (Unpublished Doctoral Dissertation). University of Manchester, UK.

Freed, B., Dewey, D., Segalowitz, N., & Halter, R. (2004). The language contact profile. *Studies in Second Language Acquisition*, *26*, 349-356.

Geeslin, K., & Gudmestad, A. (2008). The acquisition of variation in second-language Spanish: An agenda for integrating studies of the L2 sound system. *Journal of Applied Linguistics, 5*, 137-157.

Henriksen, N., Geeslin, K., & Willis, E. (2010). The development of L2 Spanish intonation during a study abroad immersion program in Leon, Spain: global contours and final boundary movements. *Studies in Hispanic and Lusophone Linguistics*, *3*(1), 113-162.

Hualde, J. I. (2005). The sounds of Spanish. Cambridge University Press.

Knouse, S. M. (2013). The acquisition of dialectal phonemes in a study abroad context: The case of the Castilian theta. *Foreign Language Annals*, *45*(4), 512-542.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the Effects of a Spanish phonetics course. *Hispania*, *88*(3), 557-567.

Lord, G. (2008). Podcasting communities and second language instruction. *Foreign Language Annals*, *41*(2), 364-379.

Lord, G. (2010). The combined effects of immersion and instruction on second language pronunciation. *Foreign Language Annals*, *43*(34), 88-503.

Preston, D. (1989). *Sociolinguistics and second language acquisition*. New York, NY: Basil Blackwell.

Preston, D. (2000). Three kinds of sociolinguistics and SLA: A psycholinguistics perspective. In F. M. B. Swierzbin, M. Anderson, C. Klee, & E. Tarone (Eds.), *Social and cognitive factors in second language acquisition: Selected Proceedings of the 1999 Second Language Research Forum* (pp. 3–30). Somerville, MA: Cascadilla.

Preston, D. (2002). A variationist perspective on SLA: Psycholinguistic concerns. In R. Kaplan (Ed.), *Oxford handbook of applied linguistics* (pp. 141–159). Oxford, England: Oxford University Press.

Salgado Robles, F. (2011). *The acquisition of sociolinguistic variation by learners of spanish in a study abroad context*. Unpublished doctoral dissertation, University of Floida, Gainesville, FL.

Segalowitz, N., & Freed, B. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, *26*, 173-199.

Simões, Antônio. (1996). Phonetics in second language acquisition: An acoustic study of fluency in adult learners of Spanish. *Hispania*, *79*, 87-95.

Stevens, John J. (2000). *The acquisition of L2 Spanish pronunciation in a study abroad context*. (Unpublished Doctoral Dissertation) Los Angeles: University of Southern California.

Rasmussen, J., & Zampini, M. (2010). The effects of phonetics training on the intelligibility and comprehensibility of native Spanish speech by second language learners. In John Levis & Kimberly LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, (pp. 38-52). Ames, IA: Iowa State University.

Real Academia Española: Database (CREA) [online]. *Corpus de referencia del español actual.* <http://www.rae.es> [April 2012].

Ringer-Hilfinger, K. (2012). Learner acquisition of dialect variation in a study abroad context: The case of the Spanish [θ]. *Foreign Language Annals*, *45*(3), 1-17.

Tarone, E. (2002). Frequency, effects, noticing, and creativity: Factors in a variationist interlanguage framework. *Studies in Second Language Acquisition, 24*, 287-296.

Willis, E., Geeslin, K., & Henriksen, N. (2009). The acquisition of /θ/ by study abroad learners in León, Spain. Paper presented at the 13th Hispanic Linguistics Symposium, San Juan, Puerto Rico.

LATE ESL LEARNERS' DIFFICULTIES OF DISTINCTION BETWEEN LAX AND TENSE VOWELS

Daniel Chang, Simon Fraser University

Calvin Weng, Simon Fraser University

The present research is to analyze the production of English lax and tense vowels by early learners of English and late learners of English (namely, Early Bilinguals and Late Bilinguals). According to Flege, Munro, and McKay's research in 1995, they discovered that Italian speakers who arrived in Canada late tended to have a noticeable foreign accent in English, whereas early Italian arrivals in Canada performed better at English pronunciation tasks. This suggests that older second language learners may have more difficulties in pronouncing native-like sounds in speech. Therefore, given that acoustic cues of vowels varied by the formant frequencies and perceived by speakers, it is likely to conclude that age of acquisition (AOA) might be the important factor that prohibits language learners to perceive, distinguish, and produce these formant frequencies correctly.

The present paper intends to analyze the speech of six Chinese Canadian English speakers differed by age of arrival in Canada. We created a reading task which contains 135 randomized one-syllable English words, purposefully mixed with lax and tense vowels. Preliminary results have shown that the Late Bilinguals are more likely to mispronounce lax and tense vowels in English than the Early Bilinguals. Therefore, the results taken from the present research have several pedagogical values, which intend to help prospective ESL educators to consider integrating pronunciation teaching in the curriculum design.

INTRODUCTION

Infants were found to have the language-general mechanism in perceiving all the sounds in the world (Best, McRoberts, & Sithole, 1988). However, by the end of the sixth month, the infants have started to pay attention to and acquired the sounds in their first language (L1), implying that they have begun to lose the ability to perceive all the sounds other than their native language (Best & MacRoberts, 2003; Werker, Gilbert, Humphrey, & Tees, 1981). Based on Lenneberg's *Critical Period Hypothesis* (1967), such inability to perceive the non-native sound contrasts was due to the loss of neural plasticity in the brain, suggesting that human beings might not achieve native-like proficiency in an L2 if the learning takes place after the critical period. According to the critical period account from Lenneberg, subsequent research has focused extensively on how age of acquisition (AOA) (Babcock, Stowe, Maloof, Brovetto, & Ullman, 2012; Best & Tyler, 2007; Birdsong, 1992, 2006; Flege, 1988, 1995; Gut, 2010; Palmer & Havelka, 2010; Piske, MacKay, & Flege, 2001; So, 2006), length of residence (LOR) (Babcock et al., 2012; Bardovi-Harlig & Bastos, 2011; Birdsong, 1992, 2006; Flege et al., 2006, 1997, 1999; Piske et al., 2001) and L1 interference (Best & Tyler, 2007; Flege et al., 2006; Picard, 2002; So, 2006) have influenced the quality of L2 learning.

Effects of Age of Acquisition (AOA) on ultimate L2 learning have been consistently reported in the literature. Studies have shown that AOA can predict the degree of foreign accent (Flege, Munro, & MacKay, 1995; Flege et al., 1999; Jia, Strange, Wu, Collado, & Guan, 2006; Piske et

al., 2001), the accuracy of morphosyntatic judgement (Birdsong, 1992, 2006; Birdsong & Molis, 2001; Johnson & Newport, 1989), and the learnability of vocabulary words (Birdsong, 1992; Izura et al., 2011; Kato, 2005; Palmer & Havelka, 2010). For example, Flege et al. (1995) discovered that early L1Italian/L2 English bilinguals in Canada tend to have more native-like English accent, whereas the late bilinguals' production of English has been judged to contain a noticeable foreign accent. This has suggested that age of acquisition of L2 is negatively correlated with the degree of perceived foreign accent and that older learners of L2 may face more difficulties in pronouncing non-native contrasts. Furthermore, Johnson & Newport (1989) examined a group of Chinese and Korean ESL learners' accuracy on grammatical judgement. Their results indicate a strong negative correlation between AOA and the accuracy of grammatical judgement. Similarly, Palmer & Havelka (2010) found out that L2 learners' speed of recognition is faster at the words which they acquired at an earlier stage. However, in Birdsong (2006)'s review of past studies on AOA and second language attainment, he has concluded that research on the effect of AOA and early bilinguals' L2 ultimate attainment were likely inconsistent due to various definitions of AOA. For instance, in a replication of Johnson & Newport (1989)'s research, after splitting the subjects into two groups with age of 16 as the cut-off line, Birdsong & Molis (2001) found out that the correlation between age and scores is not statistically significant for the early L2 learners, whereas, the relationship between the two variables is negative for the late L2 learners.

In Second Language Acquisition (SLA) research, studies have discovered that the length of stay in the target-L2-speaking country influences L2 production and perception (Bardovi-Harlig & Bastos, 2011; Best & Tyler, 2007; Flege, 1988, 1995; Gut, 2010; Jia et al., 2006; Lai, 2010; So, 2006; Sun & Heuven, 2007; Wang & Munro, 2004). For instance, in Flege et al. (2006)'s study, 155 native Korean listeners, grouped into early arrivals (6-14 years old) and late arrivals (21-38 years old), were asked to produce 8 English sentences at two different points of their residence. Study results have indicated that the late arrivals were rated as possessing a stronger foreign accent than the early arrivals; however, the early arrivals' production of English was rated lower than the native English counterparts. This findings seem to go against previous studies that report an effect of early AOA on native-like L2 learning because the finding have suggested that the early arrivals still speak English with an accent when compared to the native-born English-speaking children. Flege et al. (2006) has attributed this phenomenon to the amount of L2 exposure which the children had received compared to the adults.

Similarly, one's native language may interfere with the L2 ultimate production. Studies have shown an effect of native language on the L2 vowel production (Flege et al., 2006, 1997; Jia et al., 2006; Sun & Heuven, 2007; Wang & Munro, 2004) and prosody (Gut, 2010; Nava & Zubizarreta, 2009; Rasier & Hiligsmann, 2007; So, 2006). For instance, Jia et al. (2006) studied three groups of native Mandarin speakers with various amount of English exposure: 91 native Mandarin speakers in China and 131 native Mandarin-speaking immigrants in the United States. For the immigrant group, 131 immigrants were sub-divided into two groups with past arrivals (LOR 3-5 years) and recent arrivals (< 2 years). All participants were asked to perform an AXB discrimination task, containing pairs of English vowels, and were asked to imitate eight English vowels. Results showed that the older native Mandarin speakers in China have more accurate scores of discrimination and imitation, while the immigrant groups register no effect of age on discrimination and imitation. Their findings have also shown the influence of Mandarin phonology on the discrimination and production of English vowels, in which, for example, /ɛ/ and /æ/ were found to be the most difficult pair of discrimination. For imitation, /ʌ/ and /ɑ/ were

confusing for the Mandarin speakers as well. These results point to two important implications: one is that AOA and LOR both mutually influence L2 learning, and the other is that L2 performance is perceptually affected by the experience of L1. In the same manner, Nava & Zubizarreta (2009) have uncovered that L1 Spanish/L2 English adult learners are more likely to transfer their L1 Spanish prosody in pronouncing L2 English sentences. Results have indicated that new L1 Spanish/L2 English learners tend to preserve their L1 Spanish prosodic behaviors in reading L2 English sentences. Nevertheless, with the increase in the amount of L2 English exposure and in their L2 English proficiency, the Spanish speakers have exhibited prosodic modification, suggesting that there is a positive interference coming from Spanish to the acquisition of English prosody. As reviewed above, adult L2 learners may have difficulty achieving native-like L2 oral proficiency because research findings have revealed that adult L2 learners always speak with a noticeable accent, perceptually influenced by the phonology from their native language (Flege et al., 1995), the age that they acquired the L2 (Flege et al., 1999), or the amount of L2 exposure (Flege, 1988) in an immersion setting.

Regarding to the influence of L1 on L2 perceptual learning, two dominant theoretical accounts have been established to explain the perceptual error patterns made by L2 learners: *Perceptual Assimilation Model (PAM)* (Best, 1995) and *Speech Learning Model(SLM)* (Flege et al., 1995). PAM proposes that L2 learners tend to perceive the non-native segments based on the similarities and dissimilarities to the native L1 phonetic inventory. This model predicts that perceptual difficulty may arise from the articulatory assimilation of native and non-native segments. Secondly, SLM has attached more importance to the age-related issues in second language learning. Based on PAM, SLM predicts that the effect of age along with the experience of L1 may exert greater influences on L2 speech sound learning. That is, the spectral distance between L1 and L2 determines the formation of L2 sound categories. The difficulty to establish an L2 sound in one's mind is its similarity to the L1 phonetic inventory. If an L2 sound is perceived to be dissimilar to an L1 sound, then it will be more likely for a L2 speaker to establish a new sounding category and acquire the sound.

RATIONALE OF THE PRESENT RESEARCH AND RESEARCH QUESTIONS

As reviewed above, past studies along with these theoretical models have pointed to the implication that age of L2 acquisition (AOA) may determine the outcomes of L2 production, given that L2 learners are likely to make perceptual mistakes based on their L1 experience or the amount of L2 exposure. In a multilingual society, like Vancouver, bilingual speakers will be the group of interest in the area of SLA because upon immigration of the family, they were typically born in a bilingual setting, in which L1 and L2 exposure is simultaneous, yet L2 receives greater attention due to academic education while the L1 proficiency is still remained at a certain level. If the effect of age is salient, then it is likely that early bilingual children and late bilingual children will exhibit difference in their production of English. If this is the case, then the late bilinguals' difficulty in English production may arise from their L1 experience, in which some English vowels may be similar or dissimilar to their L1 phonetic inventory.

The following research questions are addressed, given the scope of this pilot research:

1. Are there differences in the English lax-tense vowels production among the bilinguals who differed in the age of L2 acquisition (AOA)?

2. If late L2 AOA predicts difficulty of production in L2, will L1 experience also exert negative influence on the vowel productions of L2?

METHOD

Participants

Six participants (2 males; 4 females) were selected for the present study. The mean age of the participants is 21.7 years old. They differed in the age of L2 acquisition (AOA), in which early bilinguals were defined coming to Canada before the age of 14 and having completed formal elementary education in Canada, whereas the late bilinguals coming to Canada after the age of 14 years old and completed their middle secondary education in Taiwan or in China (Flege et al., 1997). Regardless of AOA, all of our speakers were proficient bilinguals of Canadian English and Mandarin. However, proficiency level in both languages was not controlled because they were all college/university students at the point of this research experiment. It was assumed that their English proficiency should have reached a certain level in order to be admitted to the university. Table 1 shows the descriptive statistics for the six participants in the study.

Table 1

Summary of Six Participants

	Present Age	**AOA**	**Years of Schooling in Canada**
Mean	21.7	13	8.7
Standard Deviation	1.26	1.29	1.71
Range (Min; Max)	7 (19; 26)	8 (9; 17)	12 (4; 16)

Stimuli

We created a reading task stimulus that consisted of 135 randomized words. Table 1 shows the reading list that was distributed to our speakers. All the minimal papers were generated by the researchers with reference to an online English dictionary website. Five pairs of tense and lax vowel pair were categorized, such as [i-ɪ], [e-ɛ], [o-ɔ], [u-ʊ], and [a-æ] and ten random words with each vowel quality composed too, such as *sheep, meet* for */i/*, and *pit, ship* for */I/*. Twenty words per tense and lax vowel pairs were listed. In order to prevent speakers from fixating their pronunciation on recognizable monosyllabic words, twenty distracter words of /ʌ/ and bi-syllabic words were also included on the word list.

Table 2

Reading list distributed to the speakers; tense and lax vowels with distracter words

[i]	[I]	[e]	[ɛ]	[o]	[ɔ]	[u]	[ʊ]	[a]	[æ]	[ʌ]	Distracter
Bead	Kit	Cake	Egg	Old	Call	**Boot**	**Book**	Bob	Bat	but	about
Sheep	Pit	Mate	**Bet**	Boat	**Caught**	Mood	Put	Pot	Mat	Hut	Account
Peach	Pig	Date	Shed	sold	Mall	Shoot	foot	Hot	Mad	Hum	Upon
Meat	Myth	**Bait**	Met	Go	Cause	Choose	Look	Cot	Fat	Luck	Written
Deed	**Did**	Fate	Get	Told	Bought	Goose	Could	Tom	Fad	Stuff	Pause
Feed	Shit	Bake	Pet	**Coat**	Doll	Too	Good	Lot	App	Cuff	Dance
Read	Bit	Gate	Fed	Foe	Ball	Noob	Wood	Rod	Rat	Shut	But
Geek	Zit	Hate	Med	Choke	Chalk	Cool	Would	Lost	Pat	Bud	Con
Beat	Bid	Shame	Let	Mote	Fault	Soup	Cook	Boss	Hap	Dump	Stand
Sea	Sit	Nate	Set	Dote	doggy	Suit	Push	Shot	**Nat**	Study	Firm
See	Tit	Rate	Yet	Goat	Fog	Root	Butch	**Not**	Sad	Such	Fur
Cheat											Umbrella

Procedure and Data Collection

To make sure sound files are free from unnecessary noises, all of our experiments were conducted in quiet locations like library study room, or house basements. The reading task was recorded with a Logitech headset (a microphone included) connected to an ACER laptop. Wavesurfer freeware software was used as the recording device as well as the formant analysis tool for this study, because the researchers were taught to use this program in an undergraduate Phonetics training.

Upon setting up the computer and the recording device, the reading word list in paper format was placed in front of each participant. A sound testing run was performed beforehand to ensure the equipment and the microphone were fully functional. Each participant was then asked to read the word list out loud to the microphone, and was also asked to articulate clearly and slowly to avoid any possible atypical speaking rate.

Analysis

There are two stages of the analyses. On completion of the recordings gathered from the six participants, a vowel formant analysis was conducted using Wavesurfer. The built-in function within WaveSurfer's LPC and formant plot was used to determine the values of F1, F2, and F3 manually. Measurements of F1, F2, and F3 were the midpoint of the vowels, subjectively determined by the researchers. Only the underscored words shown on Table 1 were analyzed, as these words or word pairs were the minimal pairs which were commonly-seen words. Each

participant has a F1, F2, and F3 value for each word that they produced. Then, the mean scores of F1, F2, and F3 extracted from each word produced by each participant were calculated.

In order to see if there is a group difference in the production of these vowel pairs, an independent sample t-test was conducted using SPSS. The independent variable is defined as Group (Early Bilinguals, Late Bilinguals), and the dependent variables are the F1 value and the F2 value extracted from each word. F3 values are excluded because only F1 and F2 determine the height and the backness of the vowel space.

RESULTS

Figure 1 showed the overall word-by-word comparisons of formant frequencies from Early Bilinguals and Late Bilinguals. It can be generally seen that two groups produce the lax and tense vowels (e.g. [ɛ] and [e]) quite distinctively. For instance, Early Bilinguals produce the pair of "Bet" and "Bait" with distinctive F1 and F2 formant frequencies, whereas Late Bilinguals have the tendency of producing identical F1 and F2 formant frequencies. Furthermore, in the case of the front vowels (e.g. [ɛ] [e] [i] [ɪ]), Early Bilinguals have distinctive F1 and F2 values, while Early Bilinguals seem to produce close-to identical F1 values and "somehow-varied" F2 values. However, because the two figures do not really display whether a group difference exists in the production of each word with respect to F1 and F2 values, a statistical analysis was conducted using the independent-sample t test.
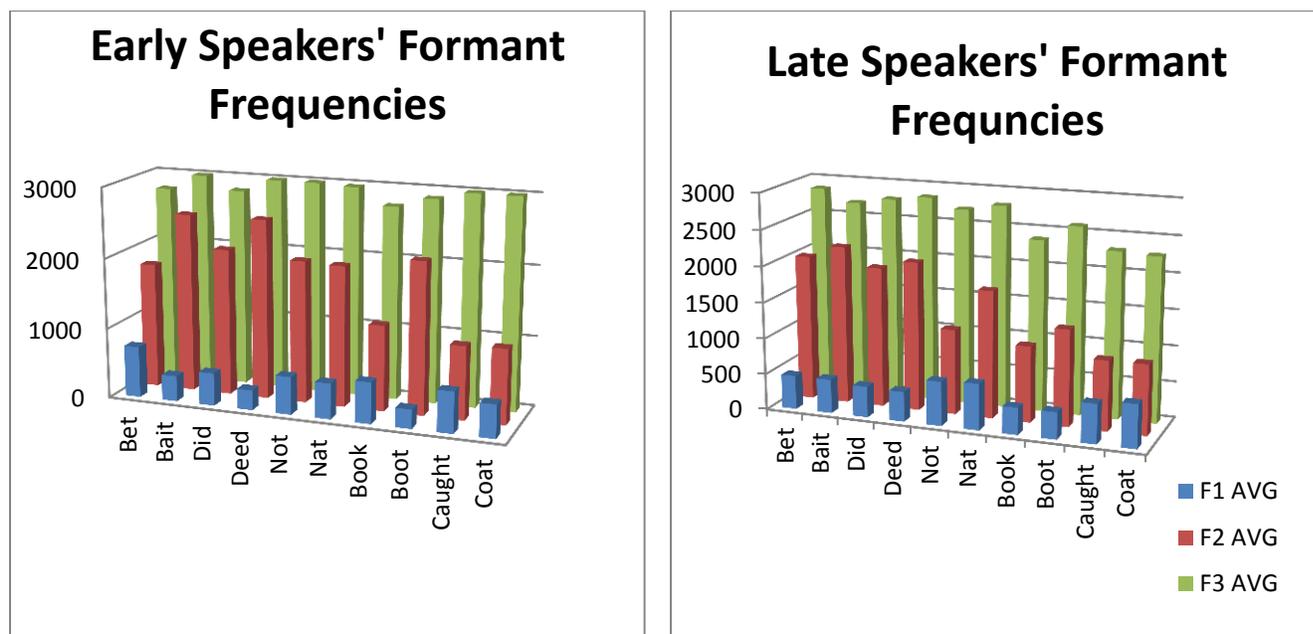
Figure 1. Overall comparison of F1, F2, and F3from early bilinguals and late bilinguals

The results from an independent-sample t-test, conducted to compare each word's vowel production by two groups (Early Bilinguals and Late Bilinguals), are displayed in Table 3 below. For the vowel [ɛ, e, ɪ, a, æ, ʊ], there was no significant difference in the F1 and F2 values for Early Bilinguals and Late Bilinguals

However, for the vowel [i], there was a significant difference in the F1 values for Early Bilinguals and Late Bilinguals; t(4)=-2.879, p=0.045 (p<0.05), but no significant difference in the F2 values for Early Bilinguals and Late Bilinguals was found; t(4)=2.193, p=0.093(p>0.05).

Furthermore, For the vowel [ɔ], although there was no significant difference in the F1 values between Early Bilinguals and Late Bilinguals for the vowel [ɔ], t(4)=0.104 , p= 0.922 (p>0.05), a significant difference in the F2 values was found, t(4)=4.867 , p= 0.008 (p<0.05). Similarly, for the vowel [o], there was a significant difference in the F1 values for Early Bilinguals and Late Bilinguals; t(4)=-3.006 , p= 0.04 (p<0.05), yet no significant difference in the F2 values for Early Bilinguals and late Bilinguals was found; t(2.1)= 1.177, p=0.358 (p>0.05).

For the vowel [u], there was no significant difference in the F1 values for Early Bilinguals and Late Bilinguals; t(4)=-1.191 , p= 0.299(p>0.05), yet a significant difference in the F2 values for Early Bilinguals and late Bilinguals was found; t(4)=3.352 , p= 0.029 (p<0.05).

Table 3

Descriptive Statistics and Independent t-test

Test Word	Group	F1 Mean (SD)	F2 Mean (SD)	t-value for F1	t-value for F2
Bet	EB	763 (135)	1873 (324)	2.423	-0.734
	LB	483 (147)	2016 (92)		
Bait	EB	340 (74)	2748 (731)	-1.524	1.439
	LB	484 (146)	2133 (104)		
Deed	EB	306 (55.8)	2698 (549)	-2.879*	2.193
	LB	413 (32.1)	1940 (238)		
Did	EB	484 (52)	2213 (444)	1.874	1.174
	LB	421 (27)	1859 (275)		
Not	EB	529 (32)	1745 (927)	-1.582	1.050
	LB	594 (63)	1179 (105)		
Nat	EB	479 (94)	1802 (667)	-2.186	0.594
	LB	614 (53)	1535 (400)		
Caught	EB	559 (94)	1049 (16)	0.104	4.867*
	LB	551 (101)	960 (27)		
Coat	EB	483 (17)	1167 (279)	-3.006*	1.177
	LB	584 (55)	976 (31)		
Boot	EB	310 (111)	2170 (25)	-1.191	3.352*
	LB	410 (94)	1444 (374)		
Book	EB	631 (144)	1255 (140)	2.219	1.983
	LB	410 (95)	1075 (72)		

*indicates p<.05; *SD indicates Standard Deviation

DISCUSSION

The main purpose of the present study is to examine whether there is an effect of early age of English acquisition on English lax-and-tense vowel productions. Although the results in the present study are not statistically significant, they are still worth further discussion, which may motivate further research on such issues with respect to bilingualism.

Production

Due to the sample size, the independent-sample t-test does not yield any significant results on the group differences of F1 and F2 values for all words, with only a few exceptions on "Deed", "Caught", "Coat", and "Boot". For instance, there is a significant group difference in the F1 values for the word "Deed", suggesting that Early Bilinguals and Late Bilinguals produce this word with a distinctive height spectral distance in the vowel space. Interestingly, for the words "Coat" and "Caught", there is a group difference in the F2 values for "Caught" and the F1 values for "Coat". Because "Caught" and "Coat" are both back vowels, it would be expected that there is no group difference in the F2 values for these two words. However, the present result shows a group difference of F2 values of "Caught", suggesting that either (1) the Late Bilinguals exhibit backness confusion for the word "Caught" or (2) the Early Bilinguals have a Canadian way of pronouncing "Caught", which is different from the way that the Late Bilinguals were taught to pronounce. The former interpretation suggests that the Late Bilinguals may have been confused with the pronunciation of "Caught" and "Coat" with respect to the spectral space. The latter interpretation, however, needs further investigation because there is no way to control for the L2 learning environment of the subjects, and lacks convincing evidence to attribute the result to a sociolinguistic factor without any qualitative data. Lastly, for the word "Boot", there is a group significance in the F2 value, suggesting that the Late Bilinguals may have been confused with the backness of the vowel [u] and [ʊ]. Therefore, this may motivate further research on the effect of L2 vowel learning experience (or L2 vowel exposure) on the ultimate L2 vowel productions. Although the present results only yield limited statistical significance, it can be seen that, from Figure 1, when these words are paired up with respect to their spectral distance, Late Bilinguals tend to be confused while producing each pair (e.g. the F1 for both "Bait" and "Bet"). Thus, consistent with previous studies that have found a younger-AOA advantage of English vowel productions (Jia et al., 2006), we anticipate that early AOA may facilitate more native-like vowel formant frequencies.

Regarding why there is a group difference in producing the lax and tense English vowels, past studies have found the influence of additional English experience on the productions of between-category vowels. From Flege et al. (1997)'s study, they have found that an additional experience with English (regardless L1 background) facilitates more native-like productions of vowel [I]. If we define the Early Bilinguals as the experienced subjects whereas the Late Bilinguals as the inexperienced ones, then we may conclude that an early additional English exposure may lead to more native-like productions of English vowels. However, we cannot make a firm conclusion at this point because of a limitation in our study where we lack of native English-speaking control groups, and do not have the vowel productions rated by native English judges.

A Possible Link between Perception and Production

Previous research has found an effect of native language background on the perception (misperception) of an L2, based on Perceptual Assimilation Model (Jia et al., 2006; Sun & Heuven, 2007). In the present study, all participants are native speakers of Mandarin, so it would be expected that they will make use of their Mandarin experience on acquiring the English vowels. From the production data, it can be seen that the Late Bilinguals tend to have difficulty of producing the target-like lax vowels with a correct spectral vowel distance. This may suggest two research areas for further research: (1) because the lax vowels in the English do not existent in Mandarin phonology, the Late Bilinguals may have the difficulty of perceiving such differences (even with extensive exposure to the L2 in the immersion setting) because they tend

to assimilate the lax-vowel category in English to the tense-vowel category in Mandarin. Therefore, the acquisition of the lax vowels in English may be unsuccessful due to the inability to perceive the differences. (2) In what period of early AOAs will it predict a more successful native-like perception of English as well as native-like production? The latter point may require a regression analysis on two factors: AOA and Perception (or Production) scores of English.

In sum, in the present study, two research questions are addressed: (1) Are there differences in the English lax-tense vowels production among the bilinguals who differed in the age of L2 acquisition (AOA)? (2) If late L2 AOA predicts difficulty of production in L2, will L1 experience also exert negative influence on the vowel productions of L2?

With respect to the former question, the descriptive statistics, outlined in Figure 1 and Table 2, have shown that there are indeed group differences between the Early Bilinguals and the Late Bilinguals' production of lax and tense vowels, but the tokens obtained from the study do not lead to a statistical analysis. Therefore, this question will be left open for further investigation. Regarding the latter question, if we regard the Early Bilinguals' productions as the baseline, the Late Bilinguals' productions of vowels are different, suggesting possible difficulty in producing the lax-and-tense vowel distinction. However, to firmly conclude such claim, a native L1-English speaking control needs to be provided.

Limitations

A major limitation from the present study is lack of the control for proficiency of the two languages in both groups, as differed proficiency in either language might have affected the L2 production. For instance, it is possible that learners with more proficient Mandarin may exhibit higher possibility of producing non-native-like English words. Furthermore, the reading list provided to the participants was chosen arbitrary; thus, it would have been ideal if a previously-established reading list has been used, so that the vowel measurements could have been yielded a better control. Lastly, the sample size in the current study is small, so the generalizability of the results is quite limited. It would have been better if future research could extend the scope of the present research in order to establish a firm conclusion. Thirdly, a rating task could have been included in order to show whether an effect of confusion occurs while the Late Bilinguals were asked to produce these vowels; furthermore, this would have allowed us to create a confusion matrix to investigate the rate of vowel confusion.

CONCLUSIONS

The present research examined whether there is an effect of early age of English acquisition on English lax-and-tense vowel productions. Descriptive statistical results have shown a possible pattern of assimilation between the lax vowels and the tense vowels by late Mandarin-English speakers. Early Mandarin-English speakers, however, may possess the advantage for producing such vowel distinctions. Since the present study only selects 5 minimal pairs of lax-and-tense vowels (as indicated in Table 2) for analysis, the rest of words will be analyzed when we record the productions from native English speakers. Therefore, these results not only motivate further research on the link between L2 perception and L2 production but also add pedagogical values for ESL teachers in pronunciation teaching.

ABOUT THE AUTHORS

Daniel Chang has a BA in Linguistics from Simon Fraser University. He continues his MA at SFU too in hoping to learn more about second language acquisition. His research interest includes Second Language Phonetics and TESL Curriculum Design.

Calvin Weng has a BA in Linguistics from Simon Fraser University. He is now continuing his TESL certification at SFU in hoping to motivate his students to learn English as their second language.

REFERENCES

Babcock, L., Stowe, J. C., Maloof, C. J., Brovetto, C. & Ullman, M. T. (2012). The storage and composition of inflected forms in adult-learned second language: A study of the influence of length of residence, age of arrival, sex, and other factors. *Bilingualism: Language and Cognition*, *15*(04), 820–840.

Bardovi-Harlig, K. & Bastos, M.-T. (2011). Proficiency, length of stay, and intensity of interaction and the acquisition of conventional expressions in L2 pragmatics. *Intercultural Pragmatics*, *8*(3), 347–384.

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp.171-204).Timonium, MD: York Press.

Best, C., McRoberts, G. W. & Sithole, N. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345–360.

Best, C. & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Best, & MacRoberts. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, *46*(2-3), 183–216.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language*, 706–755.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, *56*(s1), 9–49.

Birdsong, D. & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Manguage*, *44*(2), 235–249.

Flege, J. (1988). Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America*, *84*, 70 - 79.

Flege, J. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp.233–277). Timonium, MD: York Press

Flege, J., Birdsong, D., Bialystok, E., Mack, M., Sung, H. & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics*, *34*(2), 153–175.

Flege, J., Bohn, O.-S. & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437-470.

Flege, J., Munro, M. J. & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, *97(5)*, 3125 - 3134.

Flege, J., Yeni-Komshian, G. H. & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*(1), 78-104.

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

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marin, J., Stadthagen-González, H. & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language*, *64*(1), 32–58.

Jia, G., Strange, W., Wu, Y., Collado, J. & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America*, *119(2)*, 1118 - 1130.

Johnson, J. S. & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99.

Kato, T. (2005). Learning strategies employed by Chinese-background learners in learning Japanese vocabulary. *Japanese Studies*, *25*(3), 271–286.

Lai, Y. (2010). English vowel discrimination and assimilation by Chinese-speaking learners of English. *Concentric: Studies in Linguistics*, *36*(2), 157–182.

Lenneberg, E. H. (1967). Biological foundations of language. New York: Wiley.

Nava, E. & Zubizarreta, M. L. (2009). Order of L2 acquisition of prosodic prominence patterns: Evidence from l1 Spanish/L2 English speech. In J. Crawford, Koichi Otaki & Masahiko Takahashi (Ed.), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition (GALANA 2008)* (pp. 175–187). Somerville, MA: Cascadilla Proceedings Project

Palmer, S. D. & Havelka, J. (2010). Age of acquisition effects in vocabulary learning. *Acta Psychologica*, *135*(3), 310–315.

Picard, M. (2002). L1 interference in second language acquisition: The case of question formation in Canadian French. *IRAL*, *40*(1), 61–68.

Piske, T., MacKay, I. R. & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, *29*(2), 191–215.

Rasier, L. & Hiligsmann, P. (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique française*, *28*, 41–66.

So, C. K. L. (2006). *Effects of L1 prosodic background and AV training on learning Mandarin tones by speakers of Cantonese, Japanese, and English*. (Doctoral dissertation, Department of Linguistics-Simon Fraser University).

Sun, L. & Heuven, V. (2007). Perceptual assimilation of English vowels by Chinese listeners. Can native-language interference be predicted?. *Linguistics In the Netherlands*, 24(1), 150-161.

Wang, X. & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System*, *32*(4), 539–552.

Werker, J. F., Gilbert, J. H., Humphrey, K. & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child development*, 349–355.

Xu, W., Case, R. E. & Wang, Y. (2009). Pragmatic and grammatical competence, length of residence, and overall L2 proficiency. *System*, *37*(2), 205–216.

PROSODIC PITFALLS WHEN LEARNING SWEDISH AS A SECOND LANGUAGE

Elisabeth Zetterholm, Linnaeus University

To avoid misunderstandings in communication it is important to achieve an intelligible pronunciation, especially when learning a new language. Different factors, such as the phonology and prosody of L1 and L2, often have an influence on a second language learner's success. In Swedish it is, in particular, the fronted rounded vowels, heavy consonant clusters and the unfamiliar prosodic structure that cause problems for second language learners regardless of their native language. This paper will focus on two important prosodic features, namely vowel quantity and word stress. Recordings of Swedish language learners with different native languages have been analyzed. Even though many of them have lived in Sweden for more than 20 years they still have problems with vowel quantity and word stress. There are individual differences between the speakers, but it is also obvious that their first language might play an important role.

INTRODUCTION

New patterns in global migration affect Sweden in similar ways to many other countries. It is necessary to learn Swedish for personal and social reasons as well as to be integrated in society. Achieving an intelligible pronunciation in the new language might be a good motive to prevent a negative attitude to immigrants. In a study by Torstensson (2010) Swedes were asked about their general attitudes to immigrants and immigration. They report a positive view in general, but concerning their attitude toward different foreign accents there is a clear preference based on stereotyping and social desirability. The Swedish movie *Se upp för dårarna* (Mind the Gap) (http://www.imbd.com/title/tt0959342/) from 2007 focuses not only on a pronunciation error, but also on a general attitude to immigrants. The title is taken from one of the actors' expressions when he is working on the Metro in Stockholm and is supposed to say *Se upp för dörrarna* every time he closes the doors before departure from the station. He intended to say [seːʉpføːˌdøraɳa] (mind the gap) but he actually said [seːʉpføːˌdoːraɳa] (mind the fools). This character was a specialist in heart disease in his native country, but the only job he could get in Sweden was in the Metro. Different studies show that native speakers often judge immigrants and evaluate their personal qualities based on the speaker's foreign accent and the pronunciation plays an important role when it comes to the listener's attitude (e.g. Boyd, 2004; Flege, 1988; Munro, 2008; Torstensson, 2010). Therefore, the speaker should be motivated to acquire an intelligible pronunciation. However, it is obvious that it is difficult for some speakers to produce a near-native pronunciation depending on different factors such as transfer from their first language, differences in the languages' phonology, the speaker's age, social interaction with native speakers and motivation. It is likely that native speakers will understand a foreign-accented speech in the right context and that pronunciation errors on the segmental level are often errors of intelligibility more than of prosody (e.g. Field, 2005; Munro & Derwing, 1995; Gilakjani & Ahmadi, 2011). On the other hand, it seems that it is possible to improve intelligibility when focusing on suprasegmental features (Derwing & Rossiter, 2003). Markham (1997) shows that Swedish listeners judged speakers as native when they imitated different languages and dialects, which indicates that at least some of us can change our articulation habits. McAllistar (2000) discusses a perceptual foreign accent and its implications for L2 acquisition and reorganization

of phonetic categories, features that are missed or not correctly identified. That causes a foreign accent in both perception and production, and when it comes to practical use in e.g. simultaneous interpreting more communication training might be needed in their education. Torstensson (2010) mentions the problem of disfluencies in bilingual dialogue and indicates that the attitude to foreign accents in a courtroom may affect how immigrants are judged, which can cause problems. The study presented in this paper was done in order to shed light on pronunciation skills among some interpreters in Swedish and will only focus on the interpreters' production and the prosodic features of vowel quantity and word stress.

SWEDISH AS A SECOND LANGUAGE

In the eighties of the last century Bannert (2004) and Garlén (1988) did good work in collecting describing immigrant languages and comparing them to Swedish, partly with the implication of pointing out the potential difficulties that are likely to occur for students speaking particular native languages. These studies contain descriptions of a varied number of languages, in particular the phonological systems of the different languages, and also an analysis of observed pronunciation problems, based on a large amount of data collected from L2 learners of Swedish. In Bannert (2004) there is also a ranking of the different languages according to their distance from the phonology of the Swedish language, based on observed pronunciation errors and difficulties produced by the native speakers' diverse languages collected in the material. Some of the general problems for second language learners of Swedish are the consonant clusters, the rounded front vowels and the prosody, e.g. the vowel quantity and word stress.

Swedish vowel quantity

Swedish has nine long vowels with distinctions in both quality and quantity, /i, y, e, ø, ɛ, ʉ, u, o, ɑ/. In most dialects there are nine long and eight short vowels since there is often a coincidence between the long e and ɛ vowels in the distribution of a short vowel. Diphthongization of long vowels occurs in some dialects. The front vowels /i, y, e, ø/ appear as rounded and unrounded pairs. There is also a distinction in quality between the long and short vowels, especially for the *a* vowel (Engstrand, 2007). In the orthography the short vowel is often followed by two consonants, but that is not a rule. The quantity refers to the morphology. There are minimal pairs depending on vowel quantity, see Table 1.

Table 1
Examples of minimal pairs depending on vowel quantity.

Long vowel	Short vowel
ful [fʉːl] (ugly)	*full* [fʉl] (drunk)
glas [ɡlɑːs] (glass)	*glass* [ɡlas] (ice cream)
bus [bʉːs] (joke)	*buss* [bʊs] (bus)
här [hæːr] (here)	*herr* [hær] (mister)

Swedish word stress

There are three important prosodic distinctions to be aware of in Swedish, namely the two tonal word accents, or pitch accents, which occur on the stressed syllable in two-syllable words and make a distinction in the meaning of the words, e.g. minimal pairs. Another distinction is the

vowel quantity and the third distinction, which maybe causes most problems for second language learners since it is important for an intelligible conversation, is the word stress pattern. There are quite a number of word pairs in Swedish whose meaning is distinguishable only by word stress. However, there is a correlation between word stress and the quantity of the syllable. A stressed syllable is extended and it is only possible to have the main stress on one syllable, but more than one syllable with secondary stress. There is no fixed word stress pattern in Swedish, and the stress can occur on any of the syllables depending on the origin of the word, the inflection or whether the word is a compound word. None the less there is a simple rule saying that the first part of a compound word usually gets the main stress. Concerning inflections the rules are more complex depending on the morphology.

The stressed syllable is not marked in the orthography, which can be confusing when reading, especially if there is no clear context. One example is the word *förslag* ['fœːʂlɑːg] (the initial gesture by the conductor in front of the orchestra or the choir) or [fœʂ'lɑːg] (proposal). Another minimal pair is *förtryckt* ['fœːʈrykt] (pre-printed) or [fœ'ʈrykt] (oppressed). The rule, with some exceptions, is that some prefixes and suffixes always take stress, while others never do. See Table 2 for some examples of minimal pairs depending on stress on the first or second syllable.

Table 2.
Examples of minimal pairs depending on word stress.

Stress on first syllable	Stress on second syllable
'*dator* [dɑːtur] (computer)	*da'torer* [dɑtuːrər] (computers)
'*banan* [bɑːnan] (the track)	*ba'nan* [banɑːn] (banana)
'*kallas* [kalas] (called)	*ka'las* [kalɑːs] (party)
'*modern* [muːdəɳ] (the mother)	*mo'dern* [mudæːɳ] (fashionable)

RECORDINGS

The recordings used in this study are from oral readings of 54 sentences, which are prepared with the focus on segments and prosodic minimal pairs that are known to cause problems for learners (Bannert, 2004). Interpreters of Swedish with different mother tongues have been recorded, in order to be aware of their own pronunciation and as a training session. They were able to compare their own speech with recordings of a native speaker and they were also given individual instructions and comments. They all participated in a weekend course in Swedish Prosody for Interpreters, a further education course. Listening to their own pronunciation was one part of the course. They were told that the recordings would be used for analysis and research, and consented to this. Forty recordings of 26 female and 14 male speakers, aged 31–64, have been analyzed. Each recording used in this specific study is approximately four minutes long. The speaker's first languages vary, with a majority speaking Arabic and Kurdish, some others Albanian, Lithuanian, Polish, Russian, Serbian, Assyrian, German, Kinyarwanda/French, Pashto/Dari, Spanish or Vietnamese. The time they have lived in Sweden ranges from less than 10 years to 40 years. Most of them, but not all, have studied Swedish for Immigrants, none of them for more than two years. They are all interpreters in Swedish in different areas and with

different experiences. Some of them are licensed interpreters, while some of them have only had an introductory course. Not all of them, but some, have an education, e.g. teacher or engineer, from their native country, but find it hard to get a job in Sweden.

ANALYSES

When listening to the recordings it is obvious that some of the interpreters speak with an intelligible pronunciation and read the sentences with a clear articulation, but still with a foreign accent. However, others are hard to understand depending on the pronunciation and the way they stress the words, even though they read prepared sentences. Of course there are individual differences, but in general the deviant pronunciation seems to have a correlation to their mother tongue in the first place, not depending on how long they have lived in Sweden.

Vowel quantity

Vowel quantity is a distinctive feature in Swedish, and the recorded sentences were prepared with quite a lot of minimal pairs to get an idea of whether the speakers were aware of this prosodic feature. There are individual differences, but in general it seems that speakers with Arabic, Kurdish or Serbian as their first language do not manage to produce the contrast between long and short vowel in Swedish, regardless of how long they have lived in Sweden. An acoustic analysis in Praat (http//www.fon.hum.uva.nl/praat) confirms the audible impression. In pairs like *granen/grannen* [grɑːnən/granən] (the fir tree/the neighbor), *vägen/väggen* [vɛːgən/vɛgən] (the wall/the road) and *busar/bussar* [bʉːsar/bʉsar] (hooligans/buses) all the words have a short first vowel when produced by these speakers. Differences between speakers with other first languages are more on an individual bases. The five speakers with Albanian as their mother tongue have no problems with the length contrast or the word stress.

Word stress

Word stress is often language-specific and seems to be quite hard for second language speakers to learn and master. In a compound word the main stress is mostly on the first syllable, or at least the first word in the compound, in Swedish. More than 20 of the read sentences consist of one or two compound words. Again, most speakers with Arabic or Kurdish as their first language, seem to have problems with the Swedish way of stressing a word since they often stress the second element in the compounds. A few examples are shown in Table 3.

Table 3.
Examples of pronunciation of compound words.

Right pronunciation in Swedish	Pronounced by Arabic and Kurdish speakers
'nymålad [ˈnyːˌmoːlad] (just painted)	ny'målad [nyːˈmoːlad]
'läroböcker [ˈlæːruˌbøkər] (textbooks)	läro'böcker [læːrɔˈbøkər]
tele'fonnummer [tɛlɛˈfoːnʉmər] (phone number)	telefon'nummer [tɛlɛfoːˈnɵmər]

Some of them also have difficulties realizing the differences in the placement of word stress when pronouncing minimal pairs, especially speakers with Arabic and Kurdish as their first

languages. Figures 1–3 show some results from the acoustic analyses, made in Praat. In Figure 1 a female speaker with Arabic as her mother tongue read sentences containing the minimal pair *formel* (formula), with stress on the first syllable, and *formell* (formal), with stress on the second syllable. In both words she put the stress on the last syllable, but there should be a difference with the stress on the first syllable in the figure to the left. For comparison, Figure 2 shows the acoustic analysis of a Swedish male speaker reading the same sentences. It is obvious that there is a difference in stress and vowel quantity between the two words. In Figure 3 the vowel quantity and the word stress are compared between a male Swedish speaker (to the left) and a female native speaker of Spanish (to the right). The stress should be on the first syllable to be correct in native-like Swedish, but the Spanish speaker did not manage to do that. Her pronunciation is a nonsense word in Swedish, which might cause some problems for listeners. She has lived in Sweden for seven years, which is quite a short time compared to the other interpreters in this study.

Female speaker, L1 Arabic
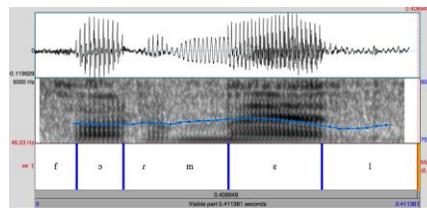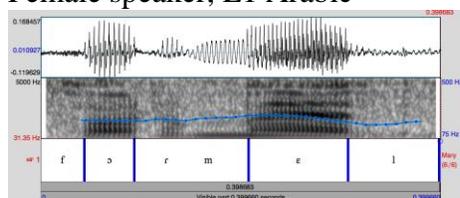
Figure 1.…en och annan formel (any formula)   *…är ganska formell* (is quite formal)

Male speaker, L1 Swedish
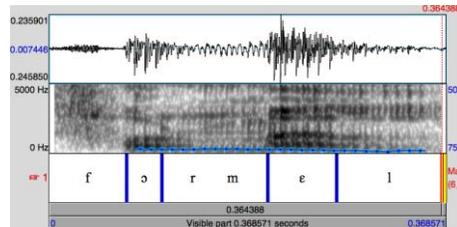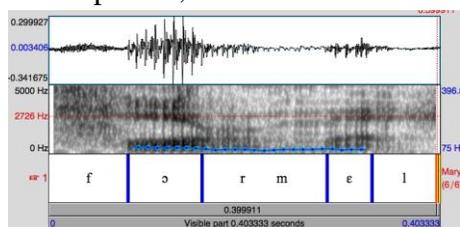
Figure 2. …en och annan formel (any formula)   *…är ganska formell* (is quite formal)
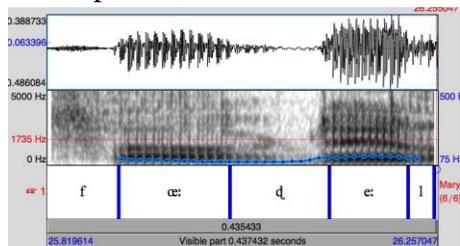
Male speaker, L1 Swedish   Female speaker, L1 Spanish
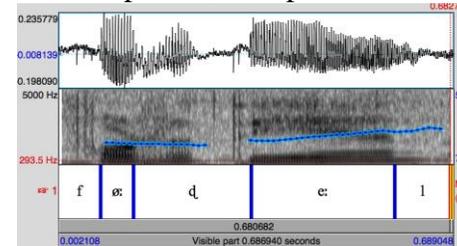
Figure 3.…en 'fördel… (an advantage)   *…en för'del* (nonsense word)

CONCLUDING REMARKS

Understandable pronunciation is important in communication in everyday life. It is sometimes frustrating for both the listener and the speaker when they do not understand each other. Both perception and production are of importance, but this paper focuses only on production. The analysis of the recordings of the interpreters in this study raises questions about their knowledge of Swedish phonology, prosody and pronunciation. There are differences between the prosody in

the Arabic and Kurdish languages compared to Swedish which can explain why these speakers have problems when producing the vowel quantity and the word stress. Referring to Bannert's ranking list (2004) Arabic seems to be one of the languages with quite a long distance from Swedish. The Kurdish language is not analyzed by Bannert. Sometimes the context is a cue for the listener, but if the pronunciation makes a change in meaning, the communication might lead to misunderstandings, which can have consequences for another person, e.g. in court or at the hospital. The results of this study also confirm earlier Swedish studies (Bannert, 2004) showing that vowel quantity and word stress are quite hard to learn for a second language learner, but necessary to control. Different research studies referred to in Field (2005) and Gilakjani and Ahmadi (2011) show that the prosody, intonation and stress patterns are often more important than an exact pronunciation of segments for understanding. A perception test with naïve Swedish listeners might confirm that in Swedish as well. The results of this study raise a couple of questions about teaching Swedish as a second language and the importance of correct pronunciation. Teachers of Swedish as a second language have to be aware of the phonetic difficulties that often have an impact on the foreign accent, both vowels and consonants, but in the light of the results of this and Bannert's (2004) study, teachers might have to give priority to the vowel quantity and stress patterns. An intelligible pronunciation probably also has a positive effect on the general attitude to different foreign accents.

ACKNOWLEDGEMENT:

ABOUT THE AUTHOR

Elisabeth Zetterholm is a senior lecturer and researcher in Swedish as a second language at Linnaeus University, Sweden. Her current research interests focus on pronunciation in second language acquisition and how to learn to speak Swedish with an intelligible accent. When teaching, most of her students are prospective teachers of Swedish as a second language and therefore she wants to develop the teaching methodology. She received her PhD in phonetics at Lund University, Sweden (2003). Her thesis and postdoctoral research were about voice imitation, with the focus on speaker identification and individual features in voice and speech.

Contact information:
Elisabeth Zetterholm, PhD
Department of Swedish
Linnaeus University
SE-351 95 Växjö
Sweden
Tel: +46 70 5354183
Email: elisabeth.zetterholm@lnu.se

REFERENCES

Bannert, R. (2004). *På väg mot svenskt uttal*. Studentlitteratur.
Boyd, S. (2004). Utländska lärare i Sverige – attityder till brytning. In: Hyltenstam, K. & Lindberg, I. (eds.), *Svenska som andraspråk – i forskning, undervisning och samhälle*. 2004:419–436. Studentlitteratur.

Derwing. T. & Rossiter, M. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. In: *Applied Language Learning*, 13 (1), pp. 1–17.

Engstrand, O. (2007). *Fonetik light*. Studentlitteratur.

Field, J. (2005). Intelligibility and the Listener: The Role of Lexical Stress. In: *TESOL Quarterly*, Vol. 39, No 3, September 2005, pp. 399–423.

Flege, J. (1988). The production and perception of speech sounds in a foreign language. In: H. Winitz (ed.), *Human communication and its disorders, a review* 1988. Norwood, NJ: Ablex, pp. 224–401.

Garlén, C. (1988). *Svenskans fonologi*. Studentlitteratur.

Gilakjani, A.P. & Ahmadi, M.R. (2011). Why Is Pronunciation So Difficult to Learn? In: *English Language Teaching*, Vol. 4, No. 3; September 2011.

Markham, D. (1997). *Phonetic Imitation, Accent, and the Learner*. Travaux de l'Institut de Linguistique de Lund 33.

McAllister, R. (2000). Perceptual Foreign Accent ant its Relevance for Simultaneous Interpreting. In: Englund Dimitrova, B. & Hyltenstam, K. (eds.), *Language Processing and Simultaneous Interpreting*. John Benjamins Publishing Company.

Munro, M.J. (2008). Foreign accent and speech intelligibility. In: Hansen Edwards J.G. & Zampini M.L. (eds.), *Phonology and Second Language Acquisition*. John Benjamins Publishing Company.

Munro, M.J. & Derwing, T.M. (1995). Foreign accent, comprehensibility and intelligibility in the speech of second language learners. In: *Language Learning*, 49 (1), pp. 285–310.

Praat: http//www.fon.hum.uva.nl/praat

Torstensson, N. (2010). *Judging the Immigrant: Accents and Attitudes*. Doctoral dissertation. Department of Language Studies, Umeå University, Sweden.

PRONUNCIATION INSTRUCTION AND SYLLABIC-PATTERN DISCRIMINATION[*]

[Rosane Silveira](), Universidade Federal de Santa Catarina

Syllable simplification strategies such as deletion (e.g., 'bag' [bæ]', 'star' [tɑɹ]) and vowel insertion (e.g., 'bag' ['bægi]', 'star' [isˈtɑɹ]) are frequent in L2 phonology. In the Brazilian Portuguese-English interphonolology, vowel insertion tends to me more recurrent than deletion. The addition of a vowel modifies the rhythm of the English language, since it creates an extra syllable, which might also result in word-stress alternation. All these changes might hinder intelligibility, and for this reason the syllable is an important curriculum component for pronunciation instruction directed at Brazilian learners of English. The present research is an investigation of the role played by pronunciation instruction in the discrimination of English CVC and CV.CV words. The participants of this study were two groups of Brazilian learners (beginners): the control group (10 students), and the experimental group (12 students). Both groups were given a discrimination pretest and posttest, between which the experimental group received instruction based on a pronunciation manual with activities focused on the English syllable and word-final consonants, whereas the control group received no such instruction. The pre and posttests consisted of an oddity discrimination test, in which the participants had to discriminate between the CVC and CV.CV words. The posttest results showed somewhat greater improvement for the experimental group than for the control group, thus suggesting a positive effect for pronunciation instruction.

INTRODUCTION

A number of researchers and educators have made a strong case for the importance of pronunciation teaching as a means of helping learners to communicate successfully (e.g., Pennington 1994; Morley, 1991 and 1994; Celce-Murcia, Goodwin and Brinton, 1996). Nevertheless, the pronunciation component has been greatly neglected in the language classroom and materials. Moreover, when this component is present in course books, it tends to be piecemeal, isolated from other language skills, and focused mostly on descriptive and controlled tasks.

As observed by Pennington (1994), the importance of pronunciation instruction lies in the fact that it can help learners to develop their interlanguage phonology by giving them "the perceptual and the productive experience they need to reconceptualize the performance targets while offering motivation to change and social experiences to develop a new value set" (p. 105).

Research has indicated that the two most important strategies of syllable simplification found in L2 renditions of impermissible syllabic patterns are consonant *deletion* and *vowel insertion* (e.g., Carlisle, 1994; Rebello, 1997, Silva Filho, 1998). When the deletion strategy is used, speakers eliminate one or more segments comprising a syllable. For instance, Vietnamese speakers of English tend to delete word-final consonants in words such as 'mine' [maɪ] (Nguyen, 1999).

When speakers resort to vowel insertion as a syllable simplification strategy, a vowel is added before or after a consonant segment (e.g., 'street' [isˈtrit], 'leg' [ˈlɛ.gi]).

Examples of how language learners resort to vowel insertion to simplify initial clusters that violate the L1 syllabic inventory can be found in the interlanguage of Brazilian and Japanese learners of English as an L2. For example, Brazilian learners may pronounce "sky" with an epenthetic vowel preceding /s/ in the /sk/ cluster: [isˈkaj] (Rebello, 1997). On the other hand, Japanese learners may pronounce the same word inserting a vowel between /s/ and /k/: [suˈkaj] (Abrahamsson, 1997). As regards word-final consonants, Brazilian learners tend to add a vowel to words ending with (a) stops, (b) some fricatives (/f, v, ʃ, ʒ/), and (c) affricates (/tʃ, dʒ/) (Silva Filho, 1998). This is illustrated by their pronunciation of words such as "tape" [ˈtejpi], "wife" [ˈwajfi], and "hush" [ˈhʌʃi].

In languages such as English, the deletion strategy is preferred, and this is attested by studies on the syllable simplification strategies employed by adult native speakers[1] when dealing with more complex consonant clusters (Temperley, 1983 and 1987; Young-Scholten & Archibald, 2000). However, in BP, vowel insertion is the strategy most commonly resorted to by adult native speakers when dealing with complex syllabic patterns in the L1 (Câmara, 1970). The preference for vowel insertion is also attested by studies investigating the acquisition of English syllabic structures by Brazilian learners (Rebello, 1997; Silva Filho, 1998). Thus, the word 'asked' is likely to be pronounced as [æst] by native speakers of English, but as [ˈæskidʒi][2] by Brazilian learners of English.

The discussion in the preceding paragraphs indicates that syllable structure is one of the areas that should be addressed by pronunciation teaching when we consider the needs of Brazilian learners of English. Although the syllable structure is widely discussed in recent theoretical phonology literature and has been a frequent item of investigation in interphonology research, it generally does not appear in pronunciation manuals or in publications on the teaching of pronunciation.

In English, all consonants except /h/ can appear in syllable-final position. Conversely, in BP only four consonants are permitted in syllable-final position: the /r/ (realized, depending on regional variety, as a trill, velar fricative, flap, or even retroflex), the lateral /l/, the nasal archiphoneme /N/, and the sibilant archiphoneme /S/ (Collischonn, 1996). However, even these are rather marginal in the coda: the /r/ tends to be deleted (*comer* "eat" [koˈme]); the /l/ is generally realized as the glide [w], or more rarely, as a dark [ɫ][3] (*mal* "bad" [maw] or [maɫ]); the /N/ loses its consonantal feature with the preceding vowel diphthongizing and assimilating the nasal feature (*bom* "good" [bõw]; leaving only the /S/ as a final consonant phonetically. Due to these constraints on syllable structure, BP speakers tend to resort to vowel epenthesis to break up

---

[1] Studies on different world languages have indicated that, in child language acquisition, the deletion strategy is also preferred over vowel insertion (Young-Scholten & Archibald, 2000).
[2] Note that Brazilians tend to pronounce the –ed morpheme as [id]. Moreover, a natural phonological process in many BP dialects is the palatalization of /d/ when it is followed by /i/.
[3] See Collischonn (1996) and Monaretto, Quednau and Hora (1996)

cross-syllabic consonant clusters in the L1. Thus, words which have not been officially modified to adapt to contemporary BP phonotactic constraints, such as *pacto* ("pact") and *advogado* ("lawyer") are pronounced with the epenthetic vowel /i/ or /e/, giving [ˈpakitu] and [adivoˈgadu] respectively.

This very productive L1 process is also known to be a frequent syllable simplification strategy in BP/English interphonology for structures such as (a) initial /s/ clusters ('stop' [isˈtɑpi]), (b) medial clusters ('substitute': [suˈbistituti], frequently with change in word-stress), (c) final clusters ('faced': [ˈfeɪsidʼ/feɪsidi], and (d) word-final singleton consonants that are not permitted in BP ('map' [ˈmæpi]. The present study investigates the effects of instruction on the perception of English word-final consonants by Brazilian learners, that is, the discrimination of CVC versus CV.CV words. Koerich (2002) has shown that learners who have difficulty producing this distinction tend to be the ones who have difficulty perceiving it; that is, they may hear a word such as 'fog' (CV<u>C</u>) as 'foggy' (CV<u>CV</u>) and vice-versa.

Some studies have shown that learners tend to build their L2 phonetic system upon the L1 system (e.g., Baptista, 2000; Flege, 1987). According to Flege (1995), language learners perceive the L2 sounds through the "filter" of the L1, which makes it difficult for them to notice certain features that are somehow different from features in similar sounds in their L1. For example, BP speakers tend to hear the English /i/ and /ɪ/ as the BP /i/, given that BP has no tense/lax contrast. Likewise, words ending in a final consonant may be heard as words ending in a vowel (e.g. 'cloud' heard as [ˈklaʊ.di]). A way of coping with this problem might be to make learners aware of the differences between the syllable structures of the two languages, as well as to show how the inappropriate transfer of L1 phonological processes (vowel insertion) can hinder communication in the L2. This new awareness should lead to better discrimination of the CVC versus CV.CV words, giving learners the opportunity to internalize the new syllable structure and, thus, ultimately produce final consonants more accurately. A combination of explicit instruction and practice of the CVC syllable structure and communicative activities using the structure might, then, be an effective way to teach the pronunciation of syllable-final consonants.

The original research tested both perception and production of the English CVC structure, but this paper reports only the perception results. Thus, assuming that better discrimination will ultimately lead to better production, the purpose of this study was to find out whether pronunciation instruction can lead to better discrimination between words ending in CVC or CV.CV.

Interphonology research on the syllable has focused on consonant clusters (e.g., Abrahamsson, 1997; Broselow, 1987; Carlisle, 1991; Cornelian Júnior, 2003; Eckman, 1991; Major, 1994; Rebello, 1997) and word-final consonants (e.g., Baptista & Silva Filho, 1997; Koerich, 2002; Silva Filho, 1998; Silveira, 2002; Yavas, 1997). A number of these studies have dealt with Brazilian learners of English and the syllable simplification strategies they resort to in order to pronounce syllabic patterns not permitted in BP, but we still need studies testing if pronunciation teaching could improve these learners' perception and production of CVC words.

The present research investigates the effect of instruction on the perception of the CVC and CV.CV words. The following research questions guided the present study:

- How do BP learners perceive the contrast between CVC words (e.g., 'fog') and CV.CV words (e.g., 'foggy')?
- How does pronunciation instruction targeting the discrimination between CVC and CV.CV words influence BP learners' perception of these contrasts?

METHOD

Participants and Instructional Content

Participants were two groups of first-semester English students from an English extension program in Brazil, a mixture of real and false beginners, many having had some English in high school. Most were graduate or undergraduate university students pursuing various majors, a few were junior high students, and some just people from the community. It was not possible to randomly select students or to match the groups for proficiency level, but the two groups were expected to be quite similar. The experimental group began with 16 students but was reduced to 12 (6 males and 6 females) with ages ranging from 18 to 28 (M = 21.83, SD = 3.01), and the control group began with 15 students and was reduced to 10 (7 males and 3 females) with ages ranging from 14 to 22 (M = 18.88, SD = 2.66). Reductions were due to the elimination of students who did not complete all the tasks of the study.

A pronunciation manual developed specifically for this study was used with the experimental group, together with the regular textbook during the instructional period. The content of the manual was limited to activities for teaching learners the differences between English and Brazilian Portuguese syllabic patterns and the inappropriateness of the use of an epenthetic vowel to overcome the articulatory problems posed by these differences. The activities developed for practice included vocabulary items with the same word-final consonants as in the discrimination test – /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /dʒ/, /m/, /n/, and /ŋ/ – but the three nasals were not practiced for lack of time.

The manual was organized according to the communicative framework suggested by Celce-Murcia et al. (1996), i.e., consisting of the following five steps: (a) description and analysis; (b) training in perception; (c) controlled practice and feedback; (d) guided practice with feedback; and (e) communicative practice and feedback.

Discrimination pretest and posttest

To test whether or not the participants could perceive the difference between monosyllabic words ending in a consonant (e.g., 'fog') and disyllabic words ending in the same consonant followed by /i/ (e.g., 'foggy'), an *oddity discrimination test* was developed, based on Flege, Munro & Fox (1994) with some adaptations. The consonants included in the discrimination test were /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /dʒ/, /m/, /n/, and /ŋ/. All word pairs were minimal pairs consisting of a CVC/CV.CV structure, with no consonant clusters and where the disyllabic word ended in /i/. A native speaker of American English recorded the sentences used in the discrimination test. The test contained sets of three sentences (Flege et al.'s (1994) version included sets of isolated syllables), where one contained a target word that differed from the other two of the same set. The carrier sentence was always "Say … now," as in the set below, where sentence "b" contains the odd item:

      a. Say *move* now.

      b. Say *movie* now.

      c. Say *move* now.

Each target consonant appeared in two of the change trials – one where the odd item out was the monosyllabic word and one where it was the disyllabic word – giving 24 target *change trials*.

As can be seen in Table 1, the test also included (a) six distracter change trials containing words dealing with other difficult vowel and consonant contrasts; (b) eight *catch trials* (as in Flege et al., 1994), where the three sentences of the set were identical: two of the catch trials contained distracters, and six of them contained the target consonant sounds /p/, /t/ and /k/. Thus, the discrimination test had a total of 38 sets of sentences; ten of the sets contained a different word in item "a", ten in item "b", ten in item "c", and eight of them (the catch trials) had no different words at all. The catch trials were expected to give some guarantee that the participants were paying attention to the three sentences of each trial, but the main analysis was based only on the change trials.

Table 1
Perception test design.

	a	b	c	The same	Total of sets	Total of sentences
Target consonants	8	8	8		24	72
Distracters	2	2	2		6	18
Catch trials				8	8	24
Total	**10**	**10**	**10**	**8**	**38**	**114**

Procedures

The data collection procedures were carried out separately for the experimental and control groups, and the discrimination and production tests were given in a single session for each. The pretest was administered to both groups in the seventh week of the course, just before the experimental group started their pronunciation instruction, and the posttest was given one week after the conclusion of the experimental group's period of pronunciation instruction.

Before beginning the discrimination test, the participants received a brief training session with three sets containing difficult vowel and consonant contrasts (similar to the distracters) to become familiar with the task. It was necessary to give the training session twice to make sure all participants understood. For both the training session and the test itself, participants were given a sheet of paper on which, for each set, they checked "a", "b", or "c" for the sentence that was different; or *todas iguais* ("all the same") if the three sentences were the same. The same procedures and materials were used in the posttest.

The English course was a 45-hour course, taught in one semester in thirty 90-minute classes meeting twice a week for 15 weeks. For the experimental group, the pronunciation lessons took up about 40 minutes of one weekly class for a period of six weeks, resulting in four hours of pronunciation instruction. These lessons were based on the manual described inthe previous subsection, and although the activities focused on pronunciation, they were also intended to be an opportunity to practice or revise the content presented in the textbook that was used as the main material in the course. During this time the control group received no pronunciation instruction regarding the features investigated, but just had more time for the normal activities in

their regular textbook. The researcher was in charge of teaching both the experimental and the control groups.

RESULTS AND DISCUSSION

As stated in the Method section, the experimental and control groups were expected to be of approximately equal proficiency. However, although the two groups obtained similar scores in the first general language skills written exam, this expectation was not borne out regarding the discrimination of CVC and CV.CV words. Table 2 shows that the control group demonstrated better discrimination of the CVC versus CV.CV words than the experimental group in the pretest change trials: 75% (M = 18; SD = 3.98) compared to 61% (M = 14.67; SD = 4.25) respectively.

As the number of participants in each group was rather small and the data presented no normal distribution, an independent sample Mann-Whitney test was run to compare the means of the pretest between groups. The results showed this difference to be significant (z = -1.88; p = .05), and a medium effect size was found (D = 0.80, r = .37). Thus, the control group was significantly better in discriminating the two syllable structures in the pretest than the experimental group.

Table 2

Frequency of correct responses in change trials of the discrimination test.

	Control Group N = 240 (10 x 24)		Experimental Group N = 288 (12 x 24)	
	Pretest	Posttest	Pretest	Posttest
No. correct	180	200	176	222
% correct	75	83	61	77
Mean per partic.	18.00	20.00	14.67	18.50
SD	3.89	3.65	4.25	5.30

Because the control group performed so much better on the pretest, the starting point of the two groups was quite different. The control group continued to perform better on the posttest: 83% (M = 20; SD = 3.65) compared to 77% (M = 18.50; SD = 5.30) respectively. However, in the posttest the difference was only 6 percentage points.

Once more a nonparametric test was used to compare the results between groups. The Mann-Whitney test showed the difference between the two groups in the posttest to be no longer significant (z = -.47, p = .63), and a small effect size was found (D = 0.32, r = .16).

The fact that the difference between the two groups was significant before treatment and was no longer significant after the treatment suggests that the treatment may have made a difference in the discrimination performance of the experimental group. Another way of verifying this tendency is by examining the gain scores from the pretest to the posttest.

The results displayed in Table 3 show that, in general, the experimental group obtained higher gain scores (M = 3.83, SD = 4.09) than the control group (M = 2; SD = 4.55). However, an independent sample Mann-Whitney test showed that, although the mean gain score for the experimental group was almost twice that of the control group, this difference between the gain scores of the two groups was not significant (z = -.64; p = .52), and a small effect size (D = .42, r = .20) was found. Thus, while it appears that pronunciation instruction might have had a slight effect on the experimental group's ability to discriminate between the CVC and CV.CV words, this hypothesis cannot be confirmed.

Table 3
Gain scores in the discrimination test from pretest to posttest.

	Control Group		Experimental Group	
Participant	Score		Participant	Score
S1	10		S11	9
S2	2		S12	3
S3	-8		S13	12
S4	1		S14	1
S5	3		S15	8
S6	1		S16	1
S7	7		S17	2
S8	0		S18	5
S9	3		S19	1
S10	1		S20	5
			S21	-2
			S22	1
Total	20			46
Mean per partic.	2.0			3.83
SD	4.55			4.09
Maximum	10			12
Minimum	-8			-2

The lack of significance in these results was probably influenced by the high standard deviations – greater than the means for both groups, which highlights the power of individual differences, a crucial factor in SLA classrooms and research. It can be seen in Table 3 that only one participant in each group (S1 and S13) achieved an increase of 10 points or more in their rates of correct responses in the discrimination posttest, and one participant in each group (S3 and S21) actually obtained worse results in the posttest. It is tempting to speculate that the somewhat higher gain scores of the experimental group may be related to the pronunciation instruction they received and that larger groups might have yielded significant results. Nevertheless, most participants in the control group also showed some improvement in their performance on the posttest, suggesting that other factors besides instruction might have influenced the posttest results, such as task familiarity and exposure to L2.

As explained in 3.2, the role of the catch trials was to verify whether the participants' responses were not mere guesses, as guessing would frequently have led to choosing an odd item out when there was none. The participants managed to correctly identify more than 80% of all catch trials in the pre and posttests. These results seem to indicate that the participants were not merely making wild guesses in the discrimination test, and thus, that the results can be assumed to faithfully represent how often the participants were actually discriminating between the pairs in the change trials.

On the other hand, there may have been an effect of the position of the odd item in each trial. There were a total of 24 change trials, where the odd item could appear in first, second, or third position. The odd targets that appeared in third ("c") position tended to trigger the lowest error rates for both experimental and control groups in the pretest, whereas for the posttest, the three

positions yielded similar rates. This result suggests a possible drawback in the design of the discrimination test, which relied heavily on the participants' ability to hold three sentences in their memories for each set and to compare them in order to identify a subtle phonological distinction. This drawback may have been less important in the posttest because of a practice effect.

FINAL CONSIDERATIONS

The results show (a) that the experimental group had significantly more difficulty with the CVC/CV.CV distinction before instruction than the control group; (b) that the experimental group continued to perform somewhat worse after treatment, but the difference between the two groups was no longer significant, thus indicating a positive effect of pronunciation instruction; and (c) that there was a tendency for greater improvement on the posttest for the experimental group, but the differences were not significant. Improvement in the posttest was found for both the experimental and the control groups, which suggests that pronunciation instruction is not the only factor influencing the acquisition of English codas, but that L2 proficiency may also play an important role. Therefore, it is possible that language exposure per se helped learners to begin to discriminate between the CVC and CV.CV words.

There were several limitations to the study, which might have interfered in the results. In addition to the small number of participants in each group, the test design itself may have caused the participants some difficulty due to the use of sentences, rather than isolated items. Memory limitations may have made those trials with the odd item in the first or second position more difficult than those where it appeared in last position. While the equal distribution of all targets among the three positions should not have biased the results in terms of which targets were more difficult, this may have made the test more difficult for those participants with lesser short-term memory capacities and thus masked improvement. This problem suggests that the use of an oddity discrimination test might not be the best way of collecting data from beginners; maybe an identification test could be used in future studies.

Although it is not possible to affirm that including the syllable in the English pronunciation syllabus led to better discrimination of English CVC/CV.CV words by Brazilian learners, there did appear to be a tendency in that direction, sufficient to encourage similar endeavors without the limitations of this exploratory study, which might achieve more conclusive results.

ABOUT THE AUTHOR

Rosane Silveira is a professor in the Department of Foreign Languages and Literature, *Universidade Federal de Santa Catarina* – Brazil. She conducts and supervises research on English-Portuguese interphonology. She was a visiting scholar at Columbia University (Teachers College) and at University of California at Los Angeles (Department of Applied Linguistics and TESL). She holds a doctoral degree from *Universidade Federal de Santa Catarina*. The author thanks the CAPES Foundation for its funding.

Contact Information:
Rosane Silveira
Universidade Federal de Santa Catarina
Rua: Joao Meirelles, 410, apt. 701
88085-200 Florianopolis, SC - Brazil
Phone number: (55) (48) 4105-0712

REFERENCES

Abrahamsson, N. (1997). Vowel epenthesis of initial /sC(C)/ cluster in Spanish speakers' L1 and L2 production: Puzzle or evidence for natural phonology? In A. James & J. Leather (Eds.), *New Sounds 97: Proceedings of the Third International Symposium on the Acquisition of Second-language Speech* (pp. 8-16). Klagenfurt, Austria: University of Klagenfurt.

Baptista, B. O. (1995). Aspectos da teoria cognitiva: Aplicações à aquisição/aprendizagem e ao ensino da pronúncia de línguas estrangeiras. *Anais do IV Congresso Brasileiro de Lingüística Aplicada*. Campinas, SP: Unicamp.

Baptista, B. O. (2000a). A pesquisa na interfonologia e o ensino da pronúncia: Procurando a interface. In M. B. M. Fortkamp & L. M. Tomitch (Eds.), *Aspectos da lingüística aplicada* (pp. 93-114). Florianópolis, Brazil: Insular.

Baptista, B. O. (2000b). *The acquisition of English vowels by Brazilian Portuguese speakers*. Advanced Research in English Series (ARES). Florianopolis: Unversidade Federal de Santa Catarina, Pós-Graduação em Inglês.

Baptista, B. O. & Silva Filho, J. (1997). The influence of markedness and syllable contact on the production of English final consonants by EFL learners. In A. James & J. Leather (Eds.), *New Sounds 97* (pp. 26-34). University of Klagenfurt, Austria.

Broselow, E. (1987). Non-obvious transfer in second language phonology. In G. Ioup & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 261-278). New York: Newbury House.

Câmara, J. M. (1970). *Estrutura da Língua Portuguesa* (8ª ed). Petrópolis, Brasil: Editora Vozes.

Carlisle, R. S. (1991). The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied Linguistics*, 12(1), 76-95.

Carlisle, R. S. (1994). Markedness and environment as internal constraints on the variability of interlanguage phonology. In H. Yavas (Ed.) *First and Second Language Phonology* (pp. 223-249). San Diego: Singular.

Celce-Murcia, M., Brinton, D., & Goodwin, J. M. (1996). *Teaching pronunciation*. New York: Cambridge University Press.

Collischonn, G. (1996). A sílaba em português. In L. Bisol (Ed.). *Introdução a estudos de fonologia do português brasileiro* (pp. 95-130). Porto Alegre, Brazil: EDIPUCRS.

Cornelian Júnior, D. (2003). *Brazilian learners' production of initial /s/ clusters: phonological structure and environment*. Unpublished master's thesis, Universidade Federal de Santa Catarina, Florianopolis, Brazil.

Eckman, F. R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of ESL learners. *Studies in Second Language Acquisition*, 13(1), 23-41.

Flege, J. E. (1987). Effects of equivalence classification on the production of foreign language speech sounds. In A. James & J. Leather (Eds.), *Sound patterns of second language acquisition* (pp. 9-39). Dordrecht: Foris.

Flege, J. E. (1995) Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.) *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). Timonium, MD: York Press.

Flege, J. E., Munro, M. J., & Fox, R. A. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustic Society of America*, 95(6), 3623-3641.

Koerich, R.D. (2002). *Perception and production of vowel epenthesis in word-final single consonant codas*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianopolis, Brazil.

Major, R. C. (1994). Chronological and stylistic aspects of second language acquisition of consonant clusters. *Language Learning*, 44(4), 655-680.

Monaretto, V. N. O., Quednau, L.R., & da Hora, D. (1996). As consoantes do português. In L. Bisol (Ed.). *Introdução a estudos de fonologia do português brasileiro* (pp. 205-246). Porto Alegre, Brazil: EDIPUCRS.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481-521.

Morley, J. (1994). A multidimensional curriculum design for speech pronunciation instruction. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New ways, new directions* (pp. 64-90). Illinois: Pentagraph Print.

Nguyen, A. (1999). The systematic reduction of English syllable final consonants. *Working Papers in Linguistics*, 6, 81-93.

Pennington, M. C. (1994). Recent research in L2 phonology: Implications for practice. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New ways, new directions* (pp. 92-107). Illinois: Pentagraph Print.

Rebello, J. T. (1997). *The acquisition of English initial /s/ clusters by Brazilian learners*. Unpublished master's thesis, Universidade Federal de Santa Catarina, Florianopolis, Brazil.

Richards, J., Hull, J., & Proctor, S. (1997). *New Interchange*: *English for international communication (book 1).* Cambridge: Cambridge University Press.

Silva Filho, J. (1998). *The production of English syllable-final consonants by Brazilian learners*. Unpublished master's thesis, Universidade Federal de Santa Catarina, Florianopolis, Brazil.

Silveira, R. (2002). The influence of pronunciation instruction on the production of English word-final consonants by Brazilian learners. *Proceedings of the Third Texas Foreign Language Education Conference (*pp. 1-30*).* Austin, TX: University of Texas at Austin.

Silveira, R. (2004). The influence of pronunciation instruction on the perception and production of English word-final consonants. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianopolis, Brazil.

Temperley, M. S. (1987). The articulatory target for final –s clusters. TESOL Quarterly, 17, 421-436.

Temperley, M. S. (1987). Linking and deletion in final consonant clusters. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 59-82). Alexandria, VA: TESOL.

Yavas, M. (1997). The effects of vowel height and place of articulation in interlanguage final stop devoicing. *International Review of Applied Linguistics*, 35(2), 115-125.

Young-Scholten, M. & Archibald, J. (2000). Second language syllable structure. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 64-101). Oxford: Basil Blackwell.

LIAISON IN L2 FRENCH: THE EFFECTS OF INSTRUCTION

Jessica L. Sturm, Purdue University

This study investigates the effects of instruction on learners' ability to make liaison in L2 French appropriately, by comparing students receiving instruction in phonetics and pronunciation ($N = 11$) with those enrolled in another advanced French course ($N = 11$). Participants were recorded reading a text aloud at the beginning and end of the semester. A group of native speakers of French ($N = 11$) were also recorded for comparison purposes. Results indicate an effect for instruction, particularly on the reduction of forbidden liaisons, and are interpreted in light of Schmidt's (1990; 1992; 1993; 1994; 1995; 2001) and Robinson's (1995; 2003) work on attention and awareness in L2 learning.

INTRODUCTION

Pronunciation is traditionally neglected in the communicative L2 classroom and in L2 acquisition studies (Hannahs, 2007). The present study looks at the effects of instruction on L2 French liaison, or the linking and resyllabification of an orthographically (but not phonetically) consonant-final word followed by a vowel-initial word in certain contexts.

Liaison often provides the cues to distinguish between singular and plural, in phrases such as *leur ami* [lœ.ʁa.mi] (their friend), as opposed to *leurs amis* [lœʁ.za.mi] (their friends), or *il aime* [i.lɛm] (he likes), versus *ils aiment* [il.zɛm] (they like). Forbidden liaison also allows for a distinction between words like *les héros* [le.e.ʁo] (the heroes) and *les zeros* [le.ze.ʁo] (the zeros).

Durand and Lyche (2008) refer to liaison as a multi-faceted phenomenon; it is obligatory, optional, or forbidden, depending on the syntactic context. Thomas (1998) mentions that liaison is rarely taught due to its complexity (p. 544), yet errors of liaison lead to serious problems. Its intricacy and its importance to nativelike speech in French prompt further investigation into how learners come to acquire it. This study contributes to the investigation by comparing learners who have been instructed in French pronunciation and phonetics for a semester with those enrolled in other advanced language courses, as well as a group of native speakers.

REVIEW OF PREVIOUS RESEARCH

L2 Pronunciation instruction

The neglect of L2 pronunciation instruction does not reduce its importance for learners. Miller (2012) noted that "the speaker needs to physically produce the sounds of the target language with enough accuracy to be understood." Derwing and Munro (2009) explain that intelligibility is crucial for communication and argue that pronunciation is learnable and that students want to learn. Harlow and Muyskens (1994) surveyed 1373 L2 learners, finding that students' number one goal is speaking, and, of fourteen stated goals, pronunciation was fifth.

Hannahs (2007) notes that there has not been a great deal of work done on L2 phonological acquisition in French (p. 51). Of those, only Thomas (2002; 2004) and Howard (2004; 2006) looked at the acquisition of liaison in L2 French. The current data are from a study based on

Lord's (2005) and Saalfeld's (2011) work with L2 Spanish learners in pronunciation courses, with learners of L2 French, focusing on at learners' use of liaison.

Liaison in French

Bybee (2001) defines liaison as "the appearance of a word-final consonant before a vowel-initial word in words that in other contexts end in a vowel." There are three categories of liaison: obligatory [il.zɛm] (*ils aiment*/they like); optional [de.za.mi.(z)ɛ̃.te.ʁe.sɑ̃] (*des amis intéressants*/interesting friends); and forbidden [ʃe//an] (*chez Anne*/at Ann's place). Liaison is difficult for learners because non-realization of a liaison is the "neutral form" of a word, according to Howard (2006). He notes that the "learner has typically learned not to produce" consonants (p. 385). Eychenne (2011) calls liaison one of the three fundamental problems of French phonology (p. 79); Thomas (2004) found that learners make three times as many liaison errors as native speakers. In short, liaison is not easy for any speaker of French, but is particularly tricky in the L2 classroom. For all the difficulties that liaison presents, Howard (2004) cautions that it cannot be considered a 'luxury' for L2 learners to learn.

One particularly difficult aspect of liaison is the role of a class of words that begin with an 'h,' which is silent in French. Words that begin orthographically with the consonant 'h' begin phonetically with a vowel sound. In one class of orthographic h-words, the unpronounced 'h' functions as a consonant, so that phenomena such as liaison are blocked: [le.a.ʁi.ko] (*les haricots*/the beans). These words are said to begin with an 'h-*aspiré*' (aspirated 'h'). The difficulty, of course, is remembering which words begin with h-*aspiré*, as there is no rule from which aspiration or non-aspiration may be predicted.

Another problematic aspect of liaison for L2 learners is that the optional liaison depends on stylistic variation; in more formal contexts, optional liaisons are made. Students presumably hear a range of speech styles from the instructors and any audio or audio-visual media presented in the classroom. Furthermore, all three types of liaison are governed by a lengthy list of rules that must be memorized (see Methodology section for details).

Attention and awareness in L2 learning

Sturm (2012) found that increased attention and awareness to an aspect of an L2 (specifically, accent marks in L2 French for L1 American English learners) that is not part of the L1 led to better retention. She based her research on Schmidt's (1990; 1992; 1993; 1994; 1995; 2001) and Robinson's (1995; 2003) concepts of attention, awareness, and noticing in L2 acquisition. Briefly stated, increased attention to new material in an L2 is believed to lead to better retention. In regards to the present study, enrollment and attendance in a course focused on pronunciation and phonetics will lead to greater awareness of and attention to targetlike use of liaison, and is hypothesized to lead to more nativelike speech in this regard.

RESEARCH QUESTION

Does instruction improve learners' use of liaison in spoken French?

METHODOLOGY

Participants (*N* = 22) were advanced undergraduate students at a large, public U.S. university. The Phonetics group (*N* = 11) were enrolled in a semester-long phonetics and pronunciation course, taught by the researcher, a native speaker of American English whose French is near-native. They were not paid for their time as the tasks involved in the study were part of the

course curriculum. The Control group ($N = 11$) were enrolled in other advanced courses but not in the phonetics and pronunciation course; they were paid for their time at the completion of the second recording. All participants were native monolingual speakers of American English, as verified by a biographical questionnaire. Two class days were devoted to liaison, and activities included listening, reading, and pronunciation exercises. The text presented liaison in reading activities, allowing students to discover rules through guided post-reading questions.

According to the course text, the following contexts for a latent word-final consonant followed by a word-initial vowel mandate a liaison: an article followed by a noun (*les amis*/the friends [le.za.mi]); an adjective followed by a noun (*bons amis*/good friends [bõ.za.mi]); a subject pronoun followed by a verb (*ils ont*/they have [il.zõ]); a subject pronoun followed by a pronoun, followed by a verb (*ils y vont*/they go there [il.zi.võ]); a verb followed by a pronoun (*allez-y*/go there [a.le.zi]); after a monosyllabic adverb (except *pas, mais,* and *trop*, when liaison is optional; *très intelligent*/very intelligent [tʁɛ.zɛ̃.te.li.ʒɑ̃]); after a monosyllabic preposition (*dans un*/ in a(n) [dɑ̃.zɛ̃]); and in certain fixed expressions.

Liaisons are forbidden, according to the course text, in the following contexts: between two syntactic groups, particularly between a subject and verb or with punctuation; between a singular noun and following adjective (*condition // adaptée*/adapted condition); before and after *et* (and); before *ou* (or); before an 'aspirated' h (h-*aspiré*); after interrogative pronouns (*Comment // est*/how is); after a pronoun following a verb (*ont-ils // un livre*/do they have a book); before a proper noun (*chez // Anne*/ at Ann's place); when a word is given as a quotation; before words beginning with the letters 'y' and 'w', except *yeux* (eyes); and in certain fixed expressions.

All other contexts in which a latent word-final consonant is followed by a word-initial vowel are considered optional liaisons and are normally pronounced in formal or careful speech. The text suggests lectures and reading aloud as situations in which optional liaisons should be made.

Participants were recorded at the beginning and end of the semester, reading a text (Appendix A), which was provided by one of the authors of the course text. According to Durand and Lyche (2008), 'reading tasks give us systematic access to much of the phonological information we seek…' (p. 38). For purposes of comparison, a group of native speakers ($N = 11$) were also recorded

For analysis, the researcher listened to the files and noted when instances of liaison occurred; Durand and Lyche (2008) also used auditory coding for their study of liaison in the Phonologie du français contemporain (PFC; Phonology of Contemporary French) corpus. The researcher considered liaison to have occurred when a latent consonant was pronounced before a word-initial vowel, whether or not the consonant had been linked with the following word.

RESULTS AND DISCUSSION

Obligatory liaison

Table 1 below illustrates the obligatory liaisons, their type, and how many were made per group/time.

Table 1

Obligatory liaisons/made per group/time

Obligatory *liaison*	Type	Natives	Phonetics Time 1	Control Time 1	Phonetics Time 2	Control Time 2
Est-il	verb+pronoun	11	10	9	10	11
est-il	verb+pronoun	11	10	10	10	10
en un	monosyllabic preposition	9	7	6	7	6
les autres	article+noun	11	11	7	10	9
autres aliments	adjective+noun	9	0	0	0	0
les enfants	article+noun	11	10	7	10	9
aux Etats	monosyllabic preposition	11	9	7	10	8
Etats-Unis	fixed expression	11	9	11	10	11
Nous aimons	subject pronoun+verb	11	11	11	11	11
nous avons	subject pronoun+verb	11	9	8	10	9
en août	monosyllabic preposition	11	9	9	9	7
dans une	monosyllabic preposition	11	11	9	11	9

Native speakers made all of the obligatory liaisons, with the exception of *autres aliments* (other foods [ɔ.tʁə.za.li.mɑ̃]), adjective+noun, and *en un* (in a/an [ɑ̃.nɛ̃]), a monosyllabic preposition. None of the learner participants in either group pronounced *autres aliments* with a liaison. The other four types of liaison included in the text are verb + pronoun (*est-il*/is it); articles (*les autres [aliments]*/the other [foods]; *les enfants*/the children); the fixed expression *Etats-Unis*/United States; and subject pronoun + verb (*nous aimons*/we like; *nous avons [bu]*/we drank). Most participants made each of these liaisons. In the phonetics group, most participants pronounced the same obligatory liaisons at both testing times (participants 1, 3, 4, 6, 7, 11). Of the other five participants, three improved by pronouncing more obligatory liaisons; while only two pronounced fewer liaisons. Among those who improved, the liaisons that were pronounced on the first recording were pronounced again on the second, as well as at least one that they had missed the first time. Participant 13 pronounced four additional obligatory liaisons. In the control group, four learners made the same number of liaisons at each recording, but not the same liaisons. For example, participant 28 pronounced nine of the twelve obligatory liaisons at each recording: at the first recording, all except *En un*, *autres aliments*, and *les enfants*; at the second,

s/he pronounced *En un* with a liaison but omitted the liaison in *aux Etats*. Only two of the control group pronounced exactly the same liaisons at both recordings. Five of the control group participants improved during the semester, while two pronounced fewer obligatory liaisons at the second recording than at the first.

Forbidden liaisons

Table 2 below illustrates the forbidden liaisons, their type, and how many were made per group/time.

Table 2

Forbidden liaisons/made per group/time

Forbidden *liaison*	Type	Natives	Phonetics, Time 1	Control, Time 1	Phonetics, Time 2	Control, Time 2
pain accompagne	across two syntactic groups	0	3	3	0	0
des haricots	h-*aspiré*	1	10	6	7	9
comment imaginer	interrogative pronoun	0	0	0	2	0
hamburgers	h-*aspiré*	1	10	7	8	8

Participants made three types of forbidden liaisons: across two syntactic groups (*pain accompagne/* bread accompanies [*pa.na.kɔ̃.paɲ]); h-*aspiré* (*des haricots*/of the beans [de.za.ri.ko] and *les hamburgers*/the hamburgers [*le.zam.byʁ.gɛʁ]); and an interrogative pronoun (*comment imaginer*/how to imagine or how could one imagine [*ko.mɑ̃.ti.ma.ʒi.ne]). Among the native speakers, only the two h-*aspiré* liaisons were made, by the same speaker. In the phonetics group, most participants pronounced the liaison on both h-*aspiré* words, although fewer learners did so on the second recording. Three participants in the phonetics group pronounced *pain accompagne* with a liaison on the first recording; none did so on the second. Only two participants (in the phonetics group) pronounced *comment imaginer* with a liaison and only on the post-test.

In the control group, three participants pronounced *pain accompagne* with a liaison, only on the pre-test. More control group participants pronounced the h-*aspiré* words with a liaison on the second recording than on the first. Taken with the phonetics group's performance, this suggests that instruction had an effect on h-*aspiré* words, as the phonetics group, as a whole, improved, while the control group regressed.

Optional liaison

As optional liaisons are, by definition, neither obligatory nor forbidden, measuring their presence or absence has no bearing on learners' intelligibility. Therefore this paper will not address pronunciation of optional liaison.

DISCUSSION

Obligatory liaison

While these results do not definitively suggest an effect of instruction, there are four phenomena worth discussion. First, two native speakers and all of the learner participants failed to make a liaison with *autres aliments.* This liaison is an adjective preceding a noun, which is a frequently heard type of liaison and often cited in the literature, e.g., *petit ami* (little friend/boyfriend [pə.ti.ta.mi]). However, *enchaînement* (resyllabification of a pronounced word-final consonant) occurs with the singular form of the adjective + noun, *autre aliment*, and participants may have thought that was the proper pronunciation for *autres aliments* as well: [*ɔt.ʁa.li.mɑ̃]. Native speakers are evidently not in agreement, as two of the eleven recorded failed to make the liaison.

Second, only nine of the eleven native speakers made a liaison with *en un*, which begins with a monosyllabic preposition. However, many number words (*onze* [11] or *huit* [eight], for example) block liaison. It may have been that the native speakers were orienting to the word *un* as the number one, rather than an article.

Third, the phonetics group, as a group, improved or stayed the same for each obligatory liaison from first to second recording. Finally, the phonetics group was more consistent in their pronunciation; nine of eleven phonetics group participants pronounced all of the obligatory liaisons they had pronounced the first time at the second recording. This was not the case for the control group, as illustrated by the discussion of participant 28.

It is tempting, but simplistic, to assume that instruction is the only reason for more consistency among the phonetics group. Saalfeld (2011) found a ceiling effect for her Spanish phonetics students in the acquisition of Spanish stress. She suggested that learners who had chosen to take an elective phonetics course were already working on their pronunciation and, consequently, were better before taking the class. The participants in the phonetics group in the current study were in the same situation as Saalfeld's (2011); they chose to take a phonetics and pronunciation class among electives in their French major and minor.

Considering Schmidt's (1990; 1992; 1993; 1994; 1995; 2001) and Robinson's (1995; 2003) work in attention, awareness, and noticing in L2 acquisition, it may also be that learners in the phonetics group were paying more attention to where they made liaisons, particularly after a semester of pronunciation instruction. In other words, instruction had an effect on the phonetics group's pronunciation of liaison by bringing liaison to students' attention. It is likely that it is both interest (indicated by enrollment in the class) and attention to the text that led to more consistent pronunciation of obligatory liaison by the phonetics group.

Additionally, it must be acknowledged that the phonetics group was more invested in doing well on the recordings, which were part of their grade. On the other hand, the control group received no benefit from doing well on either recording. As a result, they may have approached the task with a different mindset than their peers who had enrolled in the course.

Forbidden liaison

The liaisons with h-*aspiré* are especially interesting as one of the native speakers pronounced them. H-*aspiré* is difficult for non-native speakers who may be unsure which words begin with an h-*aspiré*. The topic of h-*aspiré* was covered in class and the textbook provides a selection of words beginning with the phenomenon, yet participants in the phonetics group did not improve across the semester. '*Haricots*' and '*hamburgers*' are fairly frequent words, especially in L2

language courses, where food is a popular topic, suggesting that students should know whether or not these words take a liaison. Yet most learners were erroneous in their pronunciation, at least according to the rules described in the textbook. Their overuse of liaison in this context could be ascribed to an overgeneralization of the obligatory liaison between articles and nouns.

However, one native speaker made a liaison with both h-*aspiré* words, which baffled the researcher until she was assured by a native-speaking colleague that "this is becoming the norm" (J.S. Miller, personal communication, 11 January 2012). This revelation begs the question, did learners pronounce the liaison in *des haricots* and *les hamburgers* because they didn't know not to, because they heard their instructors pronounce it, or (admittedly less likely) because they are on top of changing norms? The definitive answer is beyond the scope of this paper, but it indicates the amount of work left to be done on this topic.

Pedagogical implications

The results of this research support the call for early pronunciation/phonetics intervention in L2 classrooms. Given the difficulty presented by h-*aspiré* words, and the complexity of liaison, instruction from the beginning could help learners mentally organize when to make and not to make a liaison. Of necessity, this early pedagogical intervention would include *enchaînement* as well, as it is both part of the liaison process, and, when separate, a related phenomenon.

FINAL CONCLUSIONS

Overall, the phonetics group was more consistent in their ability to pronounce obligatory liaisons, from the beginning to the end of the semester. This does not necessarily suggest an effect of instruction, but indicates that learners in the phonetics group were more invested in their performance on the recordings. It can also be argued that learners in the phonetics group were paying more attention to their pronunciation on both recordings, particularly the second, as it was part of their final exam. As discussed by Schmidt (1990; 1992; 1993; 1994; 1995; 2001) and Robinson (1995; 2003) and as Sturm (2012) found, increased attention leads to better retention and performance in L2 learning. The phonetics group also pronounced fewer forbidden liaisons on the second recording, while the control group pronounced more. This suggests that pronunciation instruction led to better knowledge of when not to pronounce a liaison. Looking at both obligatory and forbidden liaison, the results of this study suggest that instruction leads to better use of liaison than no instruction, if for no other reason than that instruction calls learners' attention to this complex and challenging aspect of French phonetics.

ACKNOWLEDGEMENTS

The author wishes to thank Jessica Sertling Miller for reading an early draft of this paper and the anonymous reviewers for their thoughtful feedback. All errors that remain are mine.

ABOUT THE AUTHOR

Jessica L. Sturm is Assistant Professor of French and Applied Linguistics at Purdue University. She received her Ph.D. in French and Second Language Acquisition/Teacher Education from the University of Illinois at Urbana-Champaign in 2008. Her research and teaching interests include second language phonology, the use of film in the language classroom, and classroom second language acquisition, including technology in the L2 classroom.

REFERENCES

Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language and Literacy, 9*, 1, 95-114.

Borgwaldt, S.R., Hellwig, F.M., & de Groot, A.M.B. (2004) Word-initial entropy in five languages. *Written Language and Literacy, 7*(2) 165-184.

Borgwaldt, S.R., Hellwig, F.M., & de Groot, A.M.B. (2005). Onset entropy matters – Letter-to-phoneme mappings in seven languages. *Reading and Writing, 18*, 211-229.

Bybee, J. (2001). Frequency effects on French liaison. In J. L. Bybee, & P. Hopper, (Eds.) *Frequency and the emergence of linguistic structure*, (pp. 337-359). Amsterdam: John Benjamins.

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42,* 476-490.

Durand, J., & Lyche, C. (2008). French liaison in light of corpus data. *French Language Studies, 18,* 33-66.

Eychenne, J. (2011) La liaison en français et la théorie de l'optimalité. *Langue française, 169*, March, 79-101.

Hannahs, S.J. (2007). French phonology and L2 acquisition. In D. Ayoun (Ed.), *French applied linguistics*, (pp. 50-74). Amsterdam: John Benjamins.

Harlow, L.L., & Muyskens, J.A. (1994). Priorities for Intermediate-Level Language Instruction. *The Modern Language Journal, 78*, 2, 141-154.

Howard, M. (2004). Sociolinguistic variation and second language acquisition: A preliminary study of advanced learners of French. *SKY Journal of Linguistics, 17,* 143-165.

Howard, M. (2006). Variation in advanced French interlanguage: A comparison of three (socio) linguistic variables. *The Canadian Modern Language Review/La revue canadienne des langues vivantes, 62,* 3, 379-400.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, *88*, 257-267.

Miller, J.S. (2012). Teaching French pronunciation with phonetics in a college-level beginner French course. *NECTFL Review,* 69: 47-68.

Robinson, P. (1995). Attention, memory, and the 'noticing' hypothesis. *Language Learning, 45,* 2, 283-331.

Robinson, P. (2003) Attention and memory during SLA. In C.J. Doughty & M. Long (Eds.) *The Handbook of Second Language Acquisition* (pp. 631-678). Malden, MA: Blackwell Publishing.

Saalfeld, A. (2011). Acquisition of L2 phonology in advanced learners: Does instruction make a difference? In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 144-152), Ames, IA: Iowa State University.

Schmidt, R.W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129-158.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-385.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13,* 206-226.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and Awareness in Foreign Language Learning* (pp. 1-63). Honolulu: University of Hawai'i Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Sturm, J.L. (2012). Attention, awareness, and accent marks in L2 French. *Language Awareness*. Available online 18 June 2012. (print edition forthcoming).

Thomas, A. (1998). La liaison et son enseignement: Des modèles orthoépiques à la réalité linguistique. *The Canadian Modern Language Review/Revue canadienne des langues vivantes, 54,* 4, 543-552.

Thomas, A. (2002). La variation phonétique en français langue seconde au niveau universitaire avancé. *AILE, 17,* 1-14.

Thomas, A. (2004). Phonetic norm versus usage in advanced French as a second language. *IRAL, 42,* 365-382.

Appendix A. Text Recorded by Participants.

Letters involved in obligatory liaisons are **bold;** those in optional liaisons are <u>underlined</u> and those in forbidden liaisons are **<u>underlined bold</u>**

*Pourquoi le pain es**t**-il aussi important pour les Français ? Tout simplement parce qu'ils ne peuvent pa<u>s</u> envisage<u>r</u> une journée, ou même un repas, sans pain. Mais le pain es**t**-il si nécessaire pour manger ? E**n** un mo<u>t</u> oui, parce qu'il rehausse le**s** autre**s** aliment**s.***

*Pour le petit-déjeûner, rien de meilleur qu'une tartine de pai**n** avec du beurre et de la confiture. A midi, le pai**n** accompagne les salade<u>s</u> et surtout soutient le fromage. Entre les deux, il absorbe la sauce de la viande et de**s** haricots. A quatre heures, tous le**s** enfant**s** aiment manger du pain et du chocola<u>t</u> en sortant de l'école. Pour le dîner, de la soupe et du pain suffisent souvent pour un repa<u>s</u> équilibré.*

*Même au**x** Etat**s**-Unis**,** le pain fait partie de la vie. Commen<u>t</u> imaginer <u>un</u> sandwich au beurre de cacahuètes sans pain ? Et le**s** hamburgers ?*

Phrases :

*Nou**s** aimons le goût du jus que nou**s** avons bu à Honolulu e**n** août.*

*La vieille cliente achète des fruits pour Louis. Elle les mange avec lui à minuit dan**s** une ruelle.*

Translation:

Why is bread so important to the French? Simply because they cannot imagine a day, or even a meal, without bread. But is bread so necessary for eating? In a word, yes, because it enhances other foods.

For breakfast, there is nothing better than a slice of bread with butter and jam. At noon, bread accompanies salads and, above all, supports cheese. Between the two, it absorbs the sauce from the meat and beans. At four o'clock, all children love to eat bread with chocolate when leaving school. For dinner, some soup and bread often suffice for a balanced meal.

Even in the United States, bread is part of life. How could one imagine a peanut-butter sandwich without bread? Or hamburgers?

Sentences:

We like the taste of the juice we drank in Honolulu in August.

The old lady client buys fruit for Louis. She eats them with him at midnight in a passageway.

THE ROLE OF PRONUNCIATION INSTRUCTION ON THE ACQUISITION OF *LIAISONS* BY ANGLOPHONE SPEAKERS.

Nadine de Moras, Brescia University College

According to the Usage Based Model (UBM), frequency of items and structures permit young L1 learners (Ellis, 2002; Tomasello, 2003) and L2 learners (Eskildsen, 2009) to establish word properties, word sequences, and make it possible for them to master these structures. One such structure is French liaisons, where both L2 acquisition and production are of interest for several reasons. While few authors have analyzed L2 production of liaisons (Howard, 2005; Mastromonaco, 1999; Thomas, 2002), their results indicate that Anglophone students master liaisons at a near-native level. However, given the complexity of the phonetic constraints on liaisons, and the lack of information on the effect of formal instruction, I first wanted to test L2 *liaison* production. I wanted to determine whether it is possible to teach them successfully, and designed a text with 51 compulsory liaisons which 20 Francophones and 37 Anglophones read. As expected, majority Francophones pronounced obligatory liaisons in a systematic fashion (95.6%), and the three groups of Anglophone students were much less accurate (60.7 %).With the three groups of Anglophones, I tested the effect of repetitions, corrections and explanations on the acquisition of liaisons. All groups' productions increased after the intervention (average of 69.5%). The group which had the greatest progress was the one with the most repetitions.

INTRODUCTION

In French, there are two types of linking: *enchainement*, which is comparable to linking in English, and *liaison,* which is unique to French. Linking is a marker of fluent speech (Hieke, 1984) and L2 linking production affects listening comprehension (Henrichsen, 1984). Similarly, if L2 learners do not produce enchainement, they may not understand native speakers who do (Sauders, 2007).

Liaisons are important because they mark the cohesion both between two words and within a phrase. When non-native speakers do not produce the liaisons between *les* and *enfants*, for example, they separate words and may pause between the first and second words being linked (hereafter called word1 and word2). This goes against the general stress, intonation and pause patterns, and is contrary to the open syllabification of French, as well as to native speakers' expectations. The non-production of liaisons thus constitutes a possible source of confusion.

The system of linking in French is rather complex, which is why Francophone children master obligatory liaisons late, around age 6, yet continue to make mistakes in the production of the optional and so-called forbidden contexts until the age of 12 or 13 (Dugua, 2005). Adults of all ages also make mistakes with the less frequently pronounced optional and forbidden liaisons. For example, French children and adult native speakers are aware of the anti-hiatus constraint, as well as their language's open syllabification, and they go to great lengths to avoid the hiatus. Their mistakes usually consist in adding the wrong liaison consonant, or in pronouncing one which should not be pronounced.

According to the Usage Based Model (UBM), frequency of items and structures permits young L1 and L2 learners (Ellis, 2002; Gass & Mackey, 2002) to establish word properties, word sequences and to create abstract categories. Indeed, young learners of French acquire the most frequently pronounced liaisons first, and the least frequent ones last (Dugua, 2006).

Second language liaisons have seldom been studied. In Mastromonaco (1999), obligatory liaisons were reportedly produced at 93.3%; in Howard (2005, 2006), Anglophones who stayed in Ireland produced 82.5%, and students who spent a year in France produced 95% of liaisons. In Thomas (2002), Anglophones who stayed in Canada produced 93.9% of liaisons while students who spent a year in France produced 90.8%. These results (Table 1) suggest that second-language learners of French master liaisons at a near-native level.

Table 1

Summary of liaisons productions by Anglophones in Mastromonaco (1999), Thomas (2002) and Howard (2005) compared to French native speakers (Malecot, 1975).

	Malécot (1975) Francophones	Mastro-monaco	Howard Anglophones who stayed in Ireland	Howard Anglophones who went to France	Thomas Anglophones who stayed in Canada	Thomas Anglophones who stayed in Canada	Thomas Anglophones who went to France	Thomas Anglophones who went to France
Test					Pre-test	Post-test	Pre-test	Post-test
Students' level		2nd year university Canada	2nd year university Ireland	2nd year university Ireland	3rd year university Canada	3rd year university Canada	3rd year university Canada	3rd year university Canada
obligatory *liaisons* %	96.9%	**93.3%**	82.5%	**95%**	91.1	**93.9%**	86.7%	**90.8%**

However, none of these studies examined the effect of instruction. According to Thomas (1998, p. 544), this may be because it is too complex to teach.

The small amount of research on second language learning of liaisons and the reported near-native acquisition of the forms contrasts with the experience of many teachers of French. As a result, this study discusses whether Anglophone students really master liaisons at the levels reported for them. If they do not, does instruction permit progress, and if yes, to what extent? Which elements of instruction are the most beneficial for teaching French liaisons?

LIAISONS AND THE PHONETIC SYSTEM

French encourages open syllabification (syllables ending with a vowel), and favours consonant-vowel contexts and the anti-hiatus constraint (avoidance of two vowels in a row) which explains phenomena like liaisons, elisions (l'/le), epenthetic consonants (va-t-il), variable prepositions (à/en), and morpho-syntactic variation of masculine adjectives (ce/cet).

Liaison occurs when a latent consonant is pronounced and attached to the following word, if it begins with a vowel.

>   Example: *les* (word1)_(z)*enfants* (word2)   [le zã fã].

Enchainement occurs when the fixed consonant is resyllabified and pronounced with the following word, if it begins with a vowel.

> Example:	*Une bonne amie*				[yn bɔ na mi].

Both liaisons and enchainements require that word1 be resyllabified with word2, but the difference is that liaisons have a latent consonant which is otherwise not pronounced.

There are several rules for liaisons. One of them involves written <h>. There are two types of phonetic realizations of h: the mute h and the aspirated h. Even though neither h is ever pronounced in French, the two types require different rules for liaisons.

When words (of Greek or Latin origin) begin with a *mute h*, such as *habit,* elision occurs in the singular form and liaison in the plural form.

> Example:	*l'homme* (the article *le* is elided to avoid the hiatus in *le homme*) [lɔm]
>
> *les* (word1)_(z)*hommes* (word2)		[le zɔm]

When words begin with an *aspirated h*, (often of foreign origin other than Greek or Latin), such as *handicap* (from English), there is no elision and no liaison.

> Example:	*le handicap* (the article is not elided)		[lə ã di kap]
>
> *les* (word1)_(z)*handicaps* (word2)		[le ã di kap]

Liaisons are forbidden with aspirated h and with words starting with y. There are other categories of forbidden liaisons, and they can all be considered to be exceptions to the general rule that liaisons occur across word boundaries if possible. The different types of liaisons are summarized in Table 2. This classification is the most commonly used (by, among others, Delattre, 1951; Encrevé, 1988; Mastromonaco, 1999; the *Académie Française*).

Table 2
Classification of liaisons

Obligatory *liaisons*	Forbidden *liaisons*	Optional *liaisons*
Article + noun: *un_arbre*	Between a determiner and a noun starting with an aspirated h: *les hiboux*	*Liaisons* are optional if neither obligatory nor forbidden
Adjective + noun: *gros_effort*	Between a determiner and a noun starting with y: *un yoyo*	Noun + plural adjective: *Des enfants_intelligents*
After a pronoun: *nous_avons*	after *ET*: *beau et intelligent*	Auxil*iary + past participal* *Ils sont_arrivés*
After a monosylabic preposition: *en_avance*	Noun subject + verb: *Jean arrive*	Polysylabic adverb + any word: (past participal, determiner…) *beaucoup_intéressés*
Fixed expressions: *États-Unis*	Noun + singular adjective: *enfant intelligent*	negation + any word: *pas_arrivé*

In general, French final consonants are not pronounced, unless the final consonants are <c, r, f, l> or involve a liaison. These conflicting rules can potentially create confusion and hinder acquisition. Furthermore, one needs to know when and how to link the words with liaisons, and know the pronunciation rules of the liaison consonants; in addition, one must know the hierarchy of liaisons. When given a choice between pronouncing a compulsory and an optional liaison, as it is often difficult to pronounce all of them, obligatory liaisons should be pronounced first.

Experience shows that L2 learners do not seem to acquire liaisons nor enchainements naturally. They tend to separate words and produce fewer liaisons and enchainements than Francophones (Lauret, 2007, p. 59). Thus, syllabic equality, syllabification, resyllabification of French, liaisons and enchainements are difficult to master for all L2 learners of French (Charliac & Motron, 1998, pp. 7-9).

When pronouncing differently from majority Francophones, the main tendencies of non-native speakers of French (Table 3) are to either not pronounce the liaison consonant, to pronounce it with word1 instead of pronouncing it with word2, and to pronounce the final consonant improperly.

Table 3
The main ways L2 learners improperly pronounce the liaison.

Reasons of errors	erroneous pronunciation	expected pronunciation
1. Final consonant of word1 not pronounced at all, example: *Un oiseau*	[œ̃ wa zo]	[œ̃ **n**wa zo]
2. Final consonant pronounced at the end of word1, instead of at the beginning of word2 called "unlinked" *liaison* consonant, example: *ces hôpitaux*	[sɛ**z** o pi to]	[sɛ **z**o pi to]
3. Final consonant pronounced at the beginning of word2 with the wrong consonant, example: *second étage*	[s(ə) gõ de taʒ]	[s(ə) gõ **t**e taʒ]

RESEARCH QUESTIONS AND DESIGN

Research in L2 training suggests that L2 training improves students' pronunciation (Blanche, 2004; Derwing & Rossiter, 2003; Elliott, 1994; Leather, 1990; Yates 2003, etc.). Studies have generally shown improvement in pronunciation whatever the approaches and tools used. However, there are no studies on teaching French enchainements and liaisons and very little about teaching English linking (Sardegna, 2011). Sardegna's results (2011, p. 115) show significant short-term and long-improvement in accuracy while reading aloud. Our study analyzes learners' productions before and after intervention and evaluates the results of various teaching components.

Design, materials and procedure

In order to evaluate Francophones' and L2 students' productions of liaisons in certain phonetic and syntactic contexts, I designed a text (Appendix A) which included 51 obligatory liaisons, 17

optional liaisons, 14 forbidden liaisons, and 13 enchainements. The majority of the words were short and frequent.

In order to evaluate the effects of phonetic instruction, three recordings of student linking were elicited (Table 4): one before the lesson (pre-test); a second recording one week later, right after one component of the lesson (test); this second recording made it possible to analyze the effect of individual components, as the length of instruction for each element was different in each group. There was a third recording (post-test) one week after the phonetic training, once all groups had completed the three components of the lesson. The students read the same text for all three recordings.

Because the majority of the L2 participants had studied French for at least 9 years in school, they had heard and practiced the words, word sequences and syntactic structures numerous times already. The phonetic training during this experimental treatment aimed at drawing their attention to a phonetic feature (liaisons and re-syllabification) they may not have noticed or for which they may not have received explicit feedback.

Thus, I anticipated that even a short training of 30 minutes could bring some results.

The phonetic lesson had three components: **explanations** about the linking process and the concept of resyllabification; **repetitions** of sequences within the same phonetic and syntactic contexts, and **corrections**. The training focused on obligatory liaisons, and tackled briefly forbidden liaisons. There was no training with optional liaisons and enchainements.

The study took place during the lab time of a first year French course, and the lab instructors had to follow the curriculum. Moreover, it was necessary to come three times to have the students sign the forms and questionnaires, and do the three recordings (pre-test, test and post-test). Thus I only had 30 minutes to teach the lesson as I already took some time for the rest of the procedure.

In order to determine which aspects of the instruction were potentially the most useful, each group had components of the lesson in a different order, and a different amount of time was allocated to each activity. It was difficult to have corrections if students did not hear the sequences or did not have any explanations first. This is why it was difficult to analyze the role of corrections alone (Lyster, 2004; Lyster & Ranta, 2013; Lyster, Saito & Sato, 2013).

When the students did the post-test, they all showed improvement from the three components of the instruction: explanations, repetitions and corrections. However, each group had more emphasis on one of the three components. Students recorded themselves using CAN8, the system in place in the language lab. I evaluated the recordings according to a precise coding system with 12 possible codes for the pronunciation (or lack thereof) of the liaison consonants. I listened to the same recordings several times. When utterances were not clearly audible or identifiable they were discarded.

Goldvarb software was used to calculate percentages and statistics. I correlated the production of liaisons to personal information from the questionnaires and correlated the productions of liaisons with lexical and syntactic information. Goldvarb is a multivariate analysis tool used primarily in sociolinguistic variation studies (Sankoff, Tagliamonte & Smith, 2005; Tagliamonte, 2006). It determines when independent variables have a significant effect on the dependent variables, and it calculates factor weights.

Participants

I recorded 20 Francophones with a digital recorder Panasonic RR/US750, reading a text aloud:

- 12 majority Francophones: 4 from France, 1 from Belgium, 5 from Quebec, and 2 bilingual subjects having learned French from at least one parent.

- 8 Minority Francophones: 4 Franco-Ontarians and 4 Africans (Senegal, Burundi, Congo, and Mali).

I included Minority Francophones for three reasons. Firstly, I wanted to determine whether participants with more restricted contact with the French language would pronounce fewer liaisons. Secondly, Francophone teachers in school often come from Quebec, Ontario and from African countries. These Francophone teachers potentially served as the L2 learners' linguistic models as students usually hear and practice French only in class. Thus, analyzing the model could partly explain the L2 learners' results. Thirdly, many studies in French only mention native speakers from France. By having a more varied sample of Francophone speakers I thought it may represent more accurately the linguistic reality of the French-speaking communities of Canada.

The 37 Anglophone students were rated at the same intermediate level, had studied French 6-13 years (most of them at least 9 years), were aged 17-19, and were all registered in a first year French course in a Canadian university in Ontario. There were originally four groups, but in group 4, the data was complete for only two students. I ended up analyzing the overall productions of 37 students for general tendencies (whose results can be seen in Figure 1 and 2, and in Tables 5 and 6), but more specifically I looked at the 35 students in the three remaining groups.

Table 4
Procedure of the intervention for each group at Times 1, 2, 3

	Group 1 (N =10)	Group 2 (N =10)	Group 3 (N =15)
Time 1	Ethic committee form Questionnaire Explanations of procedure **PRE-TEST**	Ethic committee form Questionnaire Explanations of procedure **PRE-TEST**	Ethic committee form Questionnaire Explanations of procedure **PRE-TEST**
Time 2	1. Explanations 15 mn 2. Repetitions 8 mn 3. Corrections 5 mn 4. **TEST** 10 mn	1. Repetitions 15 mn 2. **TEST** 10 mn 3. Explanations 8 mn 4. Corrections 5 mn	1. Repetitions 10 mn 2. Corrections 10 mn 3. **TEST** 10 mn 4. Explanations 8 mn
Time 3	**POST-TEST**	**POST-TEST**	**POST-TEST**

RESULTS AND DISCUSSION

Results before phonetic training

The production of liaisons and enchainements by all participants seems to be correlated to the input which they received (Figure 1). Majority Francophones produced more obligatory liaisons (95.6%) than minority Francophones (85.5%) who in turn produced more than Anglophone students (60.7%). This result supported our initial hypothesis, namely that restricted contact with French would result in restricted productions of liaisons in comparison to majority Francophones.

Anglophones produced fewer obligatory liaisons, optional liaisons and enchainements than majority Francophones. However, Anglophone students produced more enchainements and optional liaisons than minority Francophones, which is, at first, startling. In the text they read, there were several cases of liaisons and enchainements which followed each other. It was nearly impossible to produce all of them because the reader would have to stop to take a breath; it would then be required of the reader to decide which liaison had priority. Given a choice between producing obligatory or optional liaisons, minority Francophones followed the same pattern as majority Francophones and produced obligatory liaisons first. As can be seen in Appendix B, they produced over 95% of obligatory liaisons in 10 syntactic contexts out of 12.

Anglophone students, on the contrary, did not seem to fully master the classification system. The main difficulty of the system is knowing which liaisons belong to which category, because obligatory ones should always be pronounced. The optional liaisons are a matter of style. Furthermore the production of optional liaisons can reduce that of obligatory liaisons, as it is often very difficult to pronounce all possible liaisons.

L2 learners produced a more restricted number of obligatory liaisons (60.7%) but sometimes produced optional liaisons sometimes right after an unpronounced obligatory liaison.

I believe that one can only consider a system to have been mastered if the rules are applied in a systematic way, and not solely applied to individual lexical items. L2 Anglophone students seemed not to fully master the anti-hiatus constraint or the open syllabification system of French. They had some understanding that linking exists but were not able to generalize systematically the production of liaisons within the same syntactic context, as can be seen in Appendix B.

The number of liaisons produced in this study is lower than previous research because I designed the text to test a greater variety of phonetic and syntactic contexts (other studies used an existing text with fewer possible linking contexts or tested spontaneous speech). This system of coding was stricter as to what is considered "correctly pronounced liaison consonants." Also, Francophones read the same text as Anglophones, in order to compare the same phenomenon.
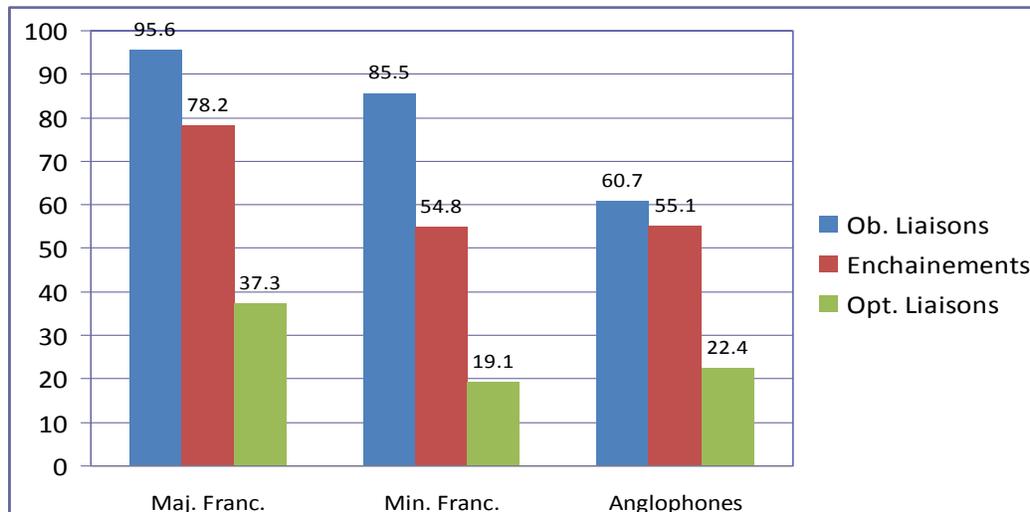
Figure 1. Productions (percentage) of obligatory and optional *liaisons* and *enchainements* by majority Francophones, minority Francophones and Anglophones at Time 1 (before the lesson).

Results after phonetic training

The next question involved the role of phonetic training. The progress following the intervention is noticeable (Figure 2). The students improved their production of obligatory and optional liaisons and enchainements at time 2 and slightly at time 3 in obligatory and optional liaisons. The fact that students produced more enchainements and optional liaisons, even though they did not receive any specific instruction in these domains, suggests a revealing rebound effect. All liaisons and enchainements are united by the same anti-hiatus constraint and open syllabification phonetic system. Students may have started to internalize the system and apply rules to word sequences which they had never heard before in all categories. This indicates that improvement may be possible after a very short amount of training.
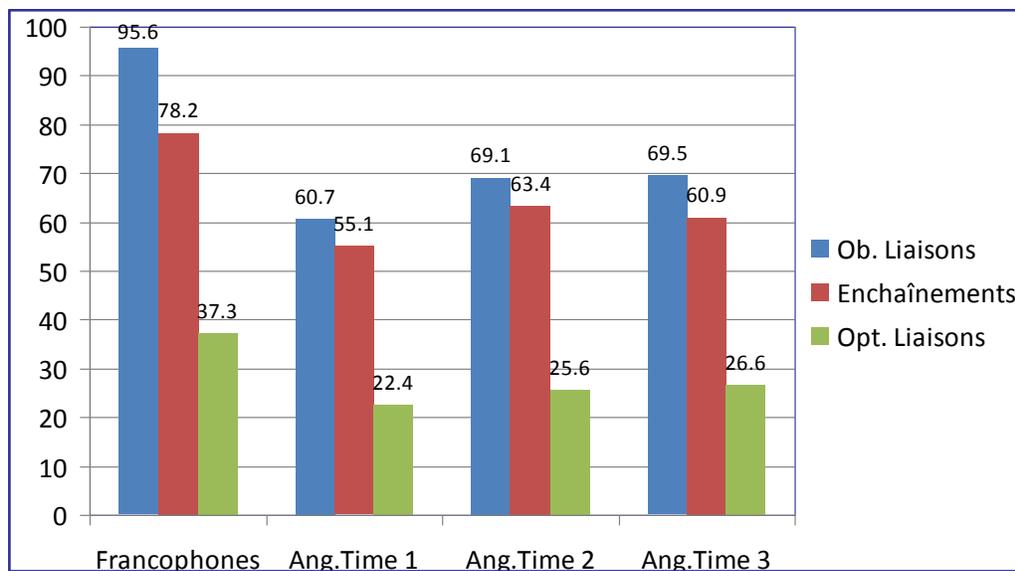

Figure 2. Productions (percentage) of obligatory and optional *liaisons* and *enchainements* by majority Francophones, and Anglophones at Time 1, Time 2 and Time 3.

The L2 learners did not pronounce the liaisons correctly in many instances because they did not attach the liaison consonant to the word2, pronouncing it at the end of word1 instead (Table 5). These liaisons mispronunciations were a subset of overall mistakes, so the percentages seem relatively low. However, the results show that Francophones virtually never pronounce unlinked consonants, except when they hesitate or when they make a pause which forces them to separate the two words (2 cases out of 1020), or to create a particular stylistic effect (Encrevé, 1988).

Table 5

*Production of unlinked **obligatory** liaisons by Francophones and Anglophones*

	Number of unlinked OL	Number of OL possible	%
All Francophones	2	1020	0.20
Anglophones time 1	167	1887	8.85
Anglophones time 2	122	1887	6.47
Anglophones time 3	126	1887	6.68

Table 6 presents overall results for obligatory liaisons for the three times with Goldvarb factor weights. A factor weight greater than 0.5 indicates liaison is favoured while a factor weight (f.w.) less than 0.5 indicates it is disfavoured. "Range" represents the differences between the highest and the lowest factor weight, multiplied by 100.

If we consider the productions of the 37 Anglophone students, we see (Table 6) an improvement at time 2 (69.1 %, f.w. = 0.529) and a slight improvement at time 3 (69.5 %, f.w. = 0.534), suggesting that phonetic training leads to quick and visible progress, which did not seem to be due to mere memorization.

Table 6

Obligatory liaisons, all participants, three times

Time of reading	*Liaison* made	N	%	factor weight
Input: .665				
Time 1	1145	1887	60.7 %	0.437
Time 2	1304	1887	69.1 %	0.529
Time 3	1311	1887	69.5 %	0.534
Range				9
TOTAL	3760	5661	66.4 %	

Results of phonetic training by groups

Even though all three groups were placed into the same class level (first year), their actual oral proficiency may have been quite different as we can see from the pre-test scores. Students had different backgrounds. Some studied French in immersion schools, others in core French, extended French or a combination of two or more of these programs. Some students had more oral practice than others, and different exposure to the language. In this university, students from different school systems are mixed in the same class level, unless they came from Francophone schools. In this case, they are often placed at a higher class level.

However, all groups seemed to improve. Group 1 (Table 7) received all components of the lesson, had the fewest repetitions (8), and the most explanations, but the group's progress was not statistically significant. Students improved slightly, but they seemed to regress the following week. This suggests that explanations may be the least efficient component of the instruction.

It may also mean that with this type of instruction, there is some regression soon after. According to Sardegna (2011, p. 115) "the higher their improvement during the course, the more they decreased in accuracy at T3." Thus, a decrease in accuracy could be part of a learning process, and not necessarily an indication of a regression in the students' learning.

Table 7
Obligatory liaisons, Anglophones, Group 1 (N=10) (explanations, corrections, repetitions)

	Liaison made	**N**	**%**	factor weight
Input: .673				
Time of reading :				
Time 1	325	510	63.7 %	
Time 2	359	510	70.4 %	ns
Time 3	346	510	67.8 %	
Range				
TOTAL	1030	1530	67.3 %	

A comparison of the productions of Group 1 (Table 7), Group 2 (Table 8) and Group 3 (Table 9) shows that Group 2, who had the most repetitions (15 minutes), made the most progress at time 2 (+11.8 %) and significant progress at time 3 (+13.4 %, w.f. = 0.552, range = 13) at the post-test. This suggests that repetitions may play a crucial role in the acquisition of correct pronunciation. It also suggests that with this type of instruction there may be less visible regression. This also indicates that weaker students may be the ones most likely to benefit from instruction.

Table 8
Obligatory liaisons, Anglophones, Group 2 (N=10) (Repetitions)

Time of reading	Liaison made	N	%	factor weight
Input: .611				
Time 1	268	510	52.5 %	0.414
Time 2	328	510	64.3 %	0.535
Time 3	336	510	65.9 %	0.552
Range				13
TOTAL	932	1530	60.9 %	

Group 3 had 10 minutes of repetitions and 10 minutes of corrections, and has the second best results after group 2 (Table 9). The students' progress is significant at time 2 (+7.4 %, f.w. = 0.525) and time 3 (+8.1 %, f.w. = 0.533). Group 1 and group 3 had a similar initial production of 63.7%. Group 3 made more progress than Group 1. Group 1 (with longer explanations) improved the least, and had the most regression at Time 3. Thus, a greater number of repetitions and corrections seem to give better results than do explanations.

The results of the three groups suggest that repetitions may give rise to better results in the short term and, possibly, in the long term as well. Since corrections also provide repetitions, they may complement repetitions while drawing attention to students' productions and encouraging them to compare their production with the target language.

Table 9
Obligatory liaisons, Anglophones, Group 3 (N=15) (Repetitions + corrections)

Time of reading	*Liaison* made	N	%	factor weight
Input: .690				
Time 1	487	765	63.7 %	0.441
Time 2	544	765	71.1 %	0.525
Time 3	549	765	71.8 %	0.533
Range				9
TOTAL	1580	2295	68.8 %	

CONCLUSION

Contrary to previous studies, Anglophone students in our study did not master liaisons at a native-like level. At T1, they produced 60.7% of obligatory liaisons (versus 95.6% for majority Francophones), 55.1% of enchainements (78.2% for majority Francophones) and 22.4% of optional liaisons (37.2% for majority Francophones).

This study shows a continuum between Majority Francophones, Minority Francophones, and L2 learners, suggesting that the greater the input received, the greater the production of liaison. This study also suggests that even a short training of 20-30 minutes can help L2 learners improve their production of liaisons and enchainements. These results are in agreement with those of other studies within the UBM framework (Eskildsen, 2009).

According to this study, repetitions play a more important role than explanations and corrections. Consequently, any linguistic activity which provides repetition may be useful for improving L2 pronunciation.

The study had several limitations. The phonetic training in this study only lasted 30 minutes and requires being replicated over a longer period of time. The post-test needs to be delayed, in order to determine the longer-term effects of instruction. When studying intermediate or advanced L2 learners, it could be useful to separate L2 participants according to their pronunciation level, established earlier by a pre-test. It is possible that the students' initial level plays a role in their progress.

ABOUT THE AUTHOR

Nadine de Moras taught ESL in France for 10 years before teaching French in Canadian universities for the past 17 years. She has been teaching advanced French language and Applied Linguistics at Brescia University College (Western University affiliate) for 6 years. Her doctoral thesis (2011) investigated the role of frequency in L1 and L2 production of *liaisons*, and the role of repetitions, explanations and corrections in L2 pronunciation training. Her research interests include L2 pronunciation acquisition and teaching, all areas of French teaching, pedagogy and material design. Publication: N. de Moras, (2010). *Guide d'écriture*. 1st Edition. Thomson et Nelson. She is currently working on a textbook on oral French, which is to be an introduction to French colloquial language.

Contact information

Address: Department of Modern Lanaguages
               Brescia University College
               1285 Western Road
               London, Ontario Canada N6G 1H2
E-mail: ndemora@uwo.ca
Telephone: 519-432-8353, ext. 28277

REFERENCES

Blanche, P. (2004). Using dictation to teach pronunciation. *Linguagem & Ensino*, 7 (1), 175-191. http://rle.ucpel.tche.br/php/edicoes/v7n1/I_patrick2.pdf

Charliac, L., & Motron, A.-C. (1999). *Phonétique progressive du français avec 600 exercices*. [Paris: CLE International.Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains.* Vermont : Middlebury College, 2e édition.

Derwing, T. M., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency and complexity of L2 accented speech. *Applied Language Learning*, 13, 1-17.

Dugua, C. (2005). De la liaison à la formation du lexique chez les jeunes enfants francophones. *Le langage et l'homme,* 40 (2), 163-182.

Dugua, C. (2008). "un nours"/"des nours" ou le rôle de la fréquence sur l'acquisition de la liaison en français. In M. Loiseau et al. (Eds.). *Autour des langues et du langage : perspective pluridisciplinaire*, Papiers sélectionnés du Colloque international des étudiants chercheurs en didactique des langues et en linguistique, Grenoble : Presses Universitaires de Grenoble. 201-208.

http://hal.archives-ouvertes.fr/docs/00/45/21/15/PDF/978-2-7061-1427-4.pdf

Elliott, A. R. (1994). Predicting near-native pronunciation in Spanish as a foreign language. *UTA. Working Papers in Linguistics*, 1, 51-68.

Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24,* 143-188.

Encrevé, P. (1988). *La liaison avec et sans enchainement. Phonologie tridimensionnelle et usages du français.* Paris: Seuil

Eskildsen, S. W. (2009). Constructing another language –Usage-Based linguistics in second language acquisition. *Applied Linguistics,* 30 (3), 335-357.

Gass, S., & Mackey, A. (2002). Frequency effects and second language acquisition: A complex picture? *Studies in Second Language Acquisition, 24(2),* 249-260.

Henrichsen, L. E. (1984). Sandhi-variation: A filter of input for learners of ESL. *Language Learning, 34*, 103-126.

Hieke, A. E. (1984). Linking as a marker of fluent speech. *Language and Speech, 27,* 243-354.

Howard, M. (2005). L'acquisition de la liaison en français langue seconde: Une analyse quantitative d'apprenants avancés en milieu guidé et en milieu naturel. *Corela*, Numéros spéciaux: Colloque AFLS.

Howard, M. (2006). Variation in advanced French interlanguage: A comparison of three (socio) linguistic variables. *The Canadian Modern Language Review/ La Revue Canadienne des Langues Vivantes*, *62 (3)*, 379-400.

Lauret, B. (2007). *Enseigner la prononciation du français: questions et outils*. Paris: Hachette.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.): *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech*. Amsterdam: University of Amsterdam, pp. 305-341.

Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching, 46* (1), 1-40.

Lyster, R., & Ranta, L. (2013). The case for variety in corrective feedback research. *Studies in Second Language Acquisition, 35* (1), 167–184.

Lyster, R. (2004). Research on form-focused instruction in immersion classrooms: Implications for theory and practice. *Journal of French Language Studies*, *14*, 321-341.

Mastromonaco, S. M. (1999). *Liaison in French as a second language*. Unpublished doctoral dissertation, University of Toronto.

Sankoff, D., S. Tagliamonte, & E. Smith. (2005). *Goldvarb X: A variable rule application for Macintosh and Windows.* Department of Linguistics, University of Toronto.

Sardegna, V. G. (2011). Pronunciation learning strategies that improve ESL learners' linking. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 105-121), Ames, IA: Iowa State University.

Sauders, G. (2007). Vowel deletion and resyllabification in French speech: a source of perceptual difficulty for American learners of French. *Phonetics Teaching & Learning Conference*. UCL August 24-26, 2007. Retrieved from http://www.phon.ucl.ac.uk/ptlc/proceedings/ptlcpaper_26e.pdf

Tagliamonte, S. (2006). *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.

Thomas, A. (1998). La liaison et son enseignement : des modèles orthoépiques à la réalité linguistique. *La Revue canadienne des langues vivantes*, *54* (4), 543-552.

Thomas, A. (2002). La variation phonétique en français langue seconde au niveau universitaire avancé. *Aile: Acquisition et Interaction en Langue Étrangère*, 17, 101-121

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Yates, K. (2003). *Teaching linguistic mimicry to improve second language pronunciation*. (Master's Thesis) Denton, Texas. UNT Digital Library. Retrieved from http://digital.library.unt.edu/ark:/67531/metadc4164/

APPENDIX A

Text read by participants with coding symbols. E = linking (*enchainement*); LO: obligatory *liaison*; LF: optional *liaison*; LI: forbidden *liaison*.

1. Cette Anglaise a demandé aux invités les affaires des étudiants.
      E1   E2   LO1   LO2   LO3

2. Tes assistants, mes ouvriers et leurs enfants sont tous nos amis.
      LO4   LO5   LF1   LO6   LO7

3. Un bon ami a dit qu'au moyen-âge, à un certain âge, on chantait en plein air.
      LO8   LO9   LO10   LF2   LO11

4. À mon avis, ton enfant ne fait aucun effort pour s'adapter à son école.
      LO12   LO13   LF3 LO14   LF4   LO15

5. Son premier amour l'a mise devant le fait accompli, ce qui est un léger ennui.
          LO16   LO17   LF5   LO18

6. Ils sont allés au dernier étage de l'ancien édifice, mais ne sont pas allés au premier.
      LF6 LF7   LO19   LO20   LF8   LF9

7. De nouveaux étudiants ont attendu au second étage du grand immeuble.
          LO21   LI1 LF10   LO22   LO23

8. Ses vieux écrits et ses nouvelles idées lui ont valu de belles acclamations.
          LO24 LF11   LO25   LO26

9. Un oiseau aux yeux bleus regarde les oies manger des yogourts dans les yachts.
      LO27   LO28   LO29   LI2   LI3

10. Ces beaux Allemands ont eu de folles aventures grâce à leurs faux-airs de gigolos.
          LO30   LI4 LF12   LO31   E3   LO32

11. Il n'y a pas de sous-entendu : on met un accent aigu sur le « e » d'États-Unis.
              LO33   LF13 LO34 LO35   LO36

12. Cet homme a parlé d'un hôpital où il y avait de vieux habits et de vieilles éponges.
      E4   E5   LO37   E6 E7   LO38 LF14   LO39

13. Des personnes ont écrit des histoires sur ces hôpitaux.
              E8 LF15   LO40   LO41

14. Les handicaps des Hollandais et des Hongrois font d'eux des héros.
      LI5   LI6   LF16   LI7   LI8

15. Deux amis de dix ans se sont vus à six heures devant les trois arbres.
          LO42   LO43   LF17 LO44   LO45

16. Il a vingt trois ans et elle a vingt cinq ans. A neuf heures, il aura cent ans.
      E9 LO46 LO47 LI9 LI10 E10 LO48 E11   LO49   E12   LO50

17. Un héros, un garçon intelligent, a eu cet accident affreux dans un bois immense.
      LI11   LI12   E13   LI13   LO51   LI14

APPENDIX B

Obligatory liaisons and grammatical categories by majority Francophones, minority Francophones and Anglophones at time 1, time 2 and time 3.

Obligatory *liaisons*	Franc.						Ang.								
	maj			min			T 1			T 2			T 3		
Grammatical structure	nb	tot	**%**	nb	tot	**%**	nb	tot	**%**	nb	tot	**%**	nb	tot	**%**
(2) definite article + noun	24	24	**100**	16	16	**100**	50	74	**67.6**	52	74	**70.3**	55	74	**74,3**
(1) demonstrative adjective + noun	12	12	**100**	8	8	**100**	26	37	**70.3**	31	37	**83.8**	30	37	**81.1**
(7) numeral adjective + noun	84	84	**100**	56	56	**100**	178	259	**68.7**	199	259	**76.8**	195	259	**75.3**
(1) preposition + noun	12	12	**100**	8	8	**100**	27	37	**73**	31	37	**83.8**	28	37	**75.7**
(5) indefinite article + noun	60	60	**100**	39	40	**97.5**	114	185	**61.6**	135	185	**73**	136	185	**73.5**
(2) indefinite adjective + noun	24	24	**100**	15	16	**93.8**	37	74	**50**	44	74	**59.5**	43	74	**58.1**
(3) preposition + article + noun	36	36	**100**	23	24	**95.8**	77	111	**69.4**	79	111	**71.2**	85	111	**76.6**
(7) possessive adjective + noun	83	84	**98.8**	56	56	**100**	222	259	**85.7**	233	259	**90**	237	259	**91.5**
(2) numeral adjective + number	23	24	**95.8**	16	16	**100**	48	74	**64.9**	54	74	**73**	65	74	**87.8**
(16) qualitative adjective + noun	182	192	**94.8**	95	128	**74.2**	289	592	**48.8**	344	592	**58.1**	338	592	**57,1**
(3) noun + adjective	28	36	**77.8**	13	24	**54.2**	56	111	**50.5**	76	111	**68.5**	74	111	**66.7**
(2) ordinal adjective + noun	17	24	**70.8**	4	16	**25**	21	74	**28.4**	26	74	**35.1**	25	74	**33.8**
Total	585	612	**95.6**	349	408	**85.5**	1145	1887	**60.7**	1304	1887	**69.1**	1311	1887	**69.5**

SCAFFOLDING STUDENTS' SELF-REGULATED EFFORTS FOR EFFECTIVE PRONUNCIATION PRACTICE

Veronica G. Sardegna, The University of Texas at Austin
Alison McGregor, The University of Texas at Austin

While attending a 15-week ITA course that empowered students with strategies to improve their English pronunciation, 15 international graduate students reflected on their use of strategies and practice at home through weekly pronunciation trackers, and self-assessed their pronunciation progress through questionnaires. To evaluate the effectiveness of student-centered instruction combined with teacher scaffolding, the students' reports on their pronunciation trackers, reflections, and self-assessments were triangulated with their accuracy scores for vowel reduction, linking, primary stress, and intonation in pre-and post-read-aloud tests. Results indicated improvement with the target features as well as evidence of effective teacher scaffolding. These findings support the view that instructors play an integral role in scaffolding students' self-regulated efforts for successful pronunciation practice.

INTRODUCTION

Research indicates that students' autonomous and self-regulated efforts are key factors contributing to their level of pronunciation improvement (He, 2011; Ingels, 2011; Sardegna, 2012). Additional factors for effective pronunciation training include explicit instruction, guided-practice, and strategy training and use (Sardegna, 2009; 2011) entwined with the critical components of learners' reflections and self-assessments (Dlaska & Krekeler, 2008; Sardegna & McGregor, 2012a). The gap in current research now lies in investigating the instructors' role in threading and integrating such factors into their curricula and activities on the basis of previous research findings. This study attempts to fill this gap in our knowledge base by examining the effect of teacher scaffolding on students' self-regulated efforts during pronunciation practice.

Teacher scaffolding consists of performing three related pedagogical actions: (a) providing a support structure to enable certain activities and skills to develop, (b) carrying out those activities in class, and (c) assisting learners in moment-to-moment interaction (Walqui, 2006). Empirical evidence documenting the positive effects of pedagogical actions undertaken in order to scaffold L2 reading, writing, listening and/or vocabulary acquisition through strategy instruction are abundant (e.g., Fitzgerald & Graves, 2004; Olson & Land, 2007; Vandergrift & Tafaghodtari, 2010; for comprehensive reviews of other studies see Chamot, 2005; Macaro, 2006). However, teacher scaffolding in relation to L2 pronunciation improvement has been an understudied area of research.

RELEVANT LITERATURE

Approaches to teaching EFL/ESL pronunciation have long emphasized the importance of empowering students with pronunciation learning strategies that they can use to correct their

mistakes (Derwing, Munro, & Wiebe, 1998; Dickerson, 1994; Morley, 1994; Osburne, 2003). Recent empirical findings on the effectiveness of pronunciation instruction continue to support metacognitive pronunciation training (He, 2011; Ingels, 2011; Sardegna, 2009; 2011; 2012), and highlight the need to consider students' beliefs about pronunciation learning and their perceptions of what works for them (Derwing & Rossiter, 2002; Couper, 2011). These views are supported by strategy training experts who have also stressed the importance of raising students' awareness as to how, when, and why strategies can be used (Chamot, 2005; Cohen & Macaro, 2007). Yet, little is known about how pronunciation teachers can facilitate awareness raising and strategy training.

Recently, Sardegna and McGregor (2012b) proposed a student-centered approach to teaching suprasegmentals grounded on research findings in the areas of pronunciation teaching, strategy instruction, and second language acquisition. The approach consists of the following components:

1. Prioritization of pronunciation goals based on student needs.

2. Empowerment through explicit instruction, guided practice, and learning strategies. That is, empowerment through teacher scaffolding.

3. Opportunities for students to monitor their performance during their pronunciation practice, and reflect on their outcomes.

Underlying this instructional model is the premise that students need to be able to recognize, understand, correct, and improve their pronunciation challenges on their own. The teacher's role is to scaffold the acquisition of these skills, thereby helping learners to become increasingly more autonomous. That is, the goal is to teach for empowerment. Under this model, students self-assess and reflect on their individual pronunciation challenges and the pronunciation training process with teacher support provided through awareness-raising activities, explicit teaching of pronunciation features and strategies, instructional guides, practice opportunities, models, resources, and individualized feedback. Yet, it is still up to the learner to utilize the opportunities, new knowledge, models, resources, and feedback to improve their pronunciation skills.

Dlaska and Krekeler (2008) alert on the importance self-assessment procedures for enhancing learners' self-awareness, increasing the reliability of their assessments, and motivating them to continue working on their L2 pronunciation problems on their own. The integral role of self-reflective practices in creating opportunities for learners to become increasingly self-regulated and more aware of their deficient oral skills has also been documented in the literature (de Saint Lèger, 2008). In fact, self-assessments have been found to be conducive to learning and more accurate when learners receive explicit instruction, feedback, and practice on how to self-assess (Chen, 2008); and when they refer to their episodic memory of using particular skills in the classroom (Ross, 1998). In accordance with these research findings, Sardegna and McGregor (2012b) argue for combining students' in- and out-of-class pronunciation practice and self-assessments through teacher scaffolding. The purpose of the current study was to look for evidence of effective teacher scaffolding in students' self-regulated efforts in a course that taught for empowerment.

Specifically, the two research questions are:

1. Do ITAs who receive English pronunciation instruction based on empowerment and their specific pronunciation needs improve their read-aloud accuracy for vowel reduction, linking, primary stress, and intonation during the course of 15 weeks?

2. What evidence of teacher scaffolding for improving these four pronunciation features is demonstrated in student-generated materials and practice behaviors out of class?

METHOD

Participants

Participants were 15 international graduate students (3 female and 12 male) from 10 different majors at a large research institution. To become eligible for teaching at this institution, international students must pass an oral English proficiency screening. Students that "conditionally pass" this screening, like the participants in this study, are required to take an oral proficiency course before they can be screened for teaching again. At the time of this study, the participants were taking this oral proficiency course. On average, their initial read-aloud accuracy level was 68% for primary stress placement, 66% for intonation, 48% for vowel reduction, and 49% for linking. Their initial accuracy level for these targets in extemporaneous speech was not measured. Their native languages were Chinese (9), Korean (4), Turkish (1), and Cantonese (1).

Characteristics of the Oral Proficiency Course

The course taught international graduate students how to employ pronunciation learning strategies to improve their English oral skills. It was specifically developed for students that wanted to become international teaching assistants (ITAs). It met for an hour and a half twice a week for fifteen weeks. The materials and activities were designed to scaffold student pronunciation learning and maximize their practice opportunities outside of class. Basically, the instructor's pedagogical intervention adhered to Sardegna and McGregor's (2012b) student-centered approach. The goal was to empower learners (i.e., to help learners become increasingly more autonomous) by instructing them on the use of pronunciation strategies and providing them with appropriate tools, resources, and materials to improve their English pronunciation. Specifically, the instructor scaffolded student learning by incorporating the following pedagogical actions in her teaching:

1. Raising students' awareness of the pronunciation features they needed to improve and the strategies they could use to improve them.

2. Providing explicit instruction and individualized feedback on specific speech features.

3. Modeling the use of resources and strategies.

4. Creating opportunities for guided practice and for learners to utilize the pronunciation strategies.

5. Developing self-assessment procedures and activities.

6. Guiding learners' reflections in relation to the learning process, the practice, and the outcomes.

These scaffolding components (or pedagogical actions) provided a support structure for the students to develop their oral skills. Before instruction, the instructor prioritized student pronunciation needs and challenges on the basis of student input, a read-aloud diagnostic test, and a three-minute academic introduction. Then, she provided explicit information about targeted speech features and pronunciation strategies for improving the features. These strategies focused on developing students' perception skills (e.g., listening for the primary stress of a sentence, identifying different meanings on the basis of changes in intonation), prediction skills (e.g., analyzing phrases in order to figure out what word(s) should receive primary stress, applying rules to judge the intonation of a phrase, dividing sentences into message units before reading them aloud), and production skills (e.g., self-monitoring, self-evaluating, reading aloud, mirroring). To maximize effectiveness, all the activities modeled guidelines and resources for practice outside of class. In addition, students received continuous feedback (recorded, written and/or oral) on homework assignments and class work, and three fifteen-minute one-on-one individualized consultations with the instructor. The consultations consisted of goal setting (including the prioritization of target speech features) and assessment of progress, and opportunities for individualized feedback. Finally, in order to make students more aware of their progress and accountable for their goals and pronunciation practice, students completed guided self-assessment and reflection activities.

Table 1 illustrates the integration of teacher scaffold into one class activity (*This I Believe: YoYoMA*) focused on improving suprasegmental features and linking. The scaffolding components or pedagogical actions embedded in this activity included teaching features and pronunciation strategies explicitly; modeling features as well as strategies to improve the features; raising awareness; offering feedback; providing resources and opportunities for practice in and out of class; and facilitating reflections and self-assessments. The activity components included a lecture on the targeted features (i.e., thought-groups, primary stress, intonation, and linking) followed by modeled and guided practice with a recorded sample script for awareness-raising, development of perception of the features, and strategy instruction. After the in-class lecture and practice with the recording and script, students were assigned to go through the remaining sentences of the script (approximately 10 sentences) for homework and mark the thought-groups and primary stresses. They were also instructed to listen specifically for intonation patterns, and linking in the same sentences. Students returned to the following class with marked scripts and discussed challenges and questions. Next, the students practiced with the marked script and recording, and tried to mimic the speech features of the model speaker (in this case, *Yo Yo Ma*). Students compared and contrasted their production with the recorded version and self-assessed their performance. If they were not satisfied with their production, they could re-record the script until they were satisfied before they submitted it for instructor feedback. Finally, to consolidate their pronunciation training experience and practice, the students reflected on the entire activity. The reflection guidelines prompted them to reflect about what they learned regarding the speech features and the strategies that can help them improve them, and what they thought of their production accuracy and obtained teacher feedback.

Table 1

This I Believe: YoYoMa Activity

Activity Components	Activity Description	Teacher Scaffolding Components
1. Lecture on suprasegmentals and linking phenomena	Instructor describes and illustrates suprasegmental features and linking.	*Giving Explicit Instruction*
2. In- and Out-of-Class Practice	Instructor provides a recorded speech[1] and a shortened transcript to practice, and teaches pronunciation strategies.	*Facilitating Learning* *Illustrating the Use of Practice Strategies*
a. Listening, Identification, Analysis, and Application of Strategies	Students listen to the recording and some pronunciation targets: • *Thought Groups*: Put a "/" where they hear a pause. • *Primary Stress*: Mark a dash over the syllables that get primary stress. • *Intonation and Rhythm*: Listen for syllable-by-syllable movements of pitch and the rhythm within thought-groups. • *Linking*: Put a "__" to indicate sounds that are linked within and across words in thought-groups.	*Raising Awareness*
b. In-Class Practice and Speech Monitoring	Teacher models targets, and students imitate and apply the strategies learned.	*Modeling Features and Strategies*
c. Out-of-Class Practice and Speech Monitoring	Students practice at home with marked script, and record own version.	*Providing Opportunities for Guided Output, Speech Monitoring and Self-Assessment*
3. Assessment	Students compare and contrast own with recorded version, and receive instructor feedback.	*Providing Opportunities for Self-Assessments and Instructor Feedback*
4. Reflection	Students write post (on Blackboard) about what they learned/practiced.	*Facilitating Guided Reflections*

Data Collection and Analysis

The effects of teacher scaffold on students' self-regulated efforts were investigated through a mixed method analysis that triangulated data from (a) a pre/post read-aloud diagnostic recording administered on the second week and again on the 15th week of class, (b) participants' background questionnaire, (c) participants' weekly pronunciation trackers, and (d) participants' self-assessments and reflections.

During the diagnostic recording, participants were told to monitor their pronunciation as best they could, but they had no way of knowing what targets were in focus for assessment and in what words or sentences in the testing materials—a long passage adapted from Celce-Murcia, Brinton, and Goodwin's (2010, p. 481) diagnostic test. The targets chosen for analysis in this study were primary stress placement, intonation, reduction, and linking. The test measured accuracy with these targets 31 times for primary stress, 25 times for intonation, 20 times for reduction in function words as well as within polysyllabic words, and 48 times for linking (including five linking types described in Sardegna, 2011). The test also measured accuracy with

---

[1] YoYoMa's *This I Believe* speech (http://www.npr.org/templates/story/story.php?storyId=87960790)

other pronunciation features, such as segmentals and word stress, which were also targeted for instruction during the course. In the post-test, students were asked to monitor their production with all the instructed pronunciation features while reading the long passage. Having received no feedback after their pre-test and with no symbols indicating a feature targeted for assessment, it is unlikely that there could be a test effect resulting from using the same test for pre- and post-instruction assessment. Participants' percentage of incorrect verbal responses was calculated by dividing the frequency of correct responses by the total number of targeted instances. The result of this calculation constituted the participant's score for each target type. Participants' percentage improvement was calculated by subtracting their pre-test mean percentage scores from their post-test mean percentage scores. These achievement scores were analyzed to measure students' reading accuracy improvement during the 15-week intervention. No analyses were performed to measure students' improvement accuracy in extemporaneous speech.

The background questionnaire administered at the beginning of instruction elicited information about participants' perceptions of their pronunciation problems. During the course, students were free to choose type, amount, and frequency of practice activities outside of class. The weekly pronunciation trackers were designed to gather participants' choice of resources and activities as well as the amount of time and frequency of practice. For this study, the researchers only analyzed the type of activities participants self-selected for practice. The guided self-assessments and reflections elicited participants' experiences with the features and activities, and their conclusions on what they thought they needed to do to improve the features. They also asked them to reflect on the value and effectiveness of the activities they performed. Students' self-reported data on the background questionnaire, weekly pronunciation trackers, and self-assessments were analyzed for evidence of teacher scaffold effectiveness in guiding students' self-regulated efforts for pronunciation improvement.

RESULTS

Research Question 1: Do ITAs who receive pronunciation instruction based on empowerment and on their specific pronunciation needs improve their read-aloud accuracy for vowel reduction, linking, primary stress, and intonation during the course of 15 weeks?

Participants' achievement scores from the pre/post read-aloud diagnostic recorded on the second week and again on the 15th week of class were compared and analyzed for significant differences through paired t-tests. When students' overall accuracy for these features was compared, a paired t-test revealed that their accuracy improved 11.5% ($SD = 7.574$; $p = 0.000$). A series of paired t-tests were performed to analyze for significant improvement with each of the targeted features. These tests revealed that the students improved each of the features significantly ($p < 0.05$). Specifically, students' improvement with reading accuracy in reduction was 5.7%, with linking was 12.9%, with primary stress was 14.5%, and with intonation was 11.4% (see Figure 1).
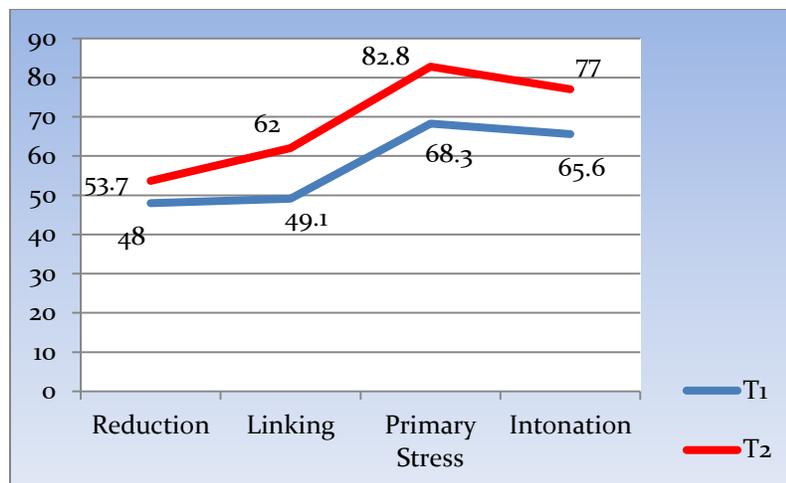
Figure 1. Reading accuracy improvement with target features from T1 (before instruction) to T2 (after instruction).

It was therefore concluded that the ITAs who received scaffolded instruction based on empowerment and on their specific pronunciation needs improved their read-aloud accuracy with reduction, linking, primary stress, and intonation targets during the course of 15 weeks. Because there was no measurement of extemporaneous speech accuracy, no claims are made regarding students' improvement with these features in spontaneous speech.

Research Question 2: What evidence of teacher scaffolding with these features is demonstrated in student-generated materials and practice behaviors out of class?

The background questionnaire that students completed before instruction revealed that half of the participants (n=8) believed that their pronunciation problems were related to segmental problems, while a few identified problems with words (n=3) or sentences in general (n=1), or reported not knowing what their problems were (n=3). Interestingly, only S8 and S15 identified problems with suprasegmental features in addition to problems with segmental features. In summary, students initially self-assessed their pronunciation challenges as difficulty in pronouncing specific sounds or words, or were unsure of their problems.

After an awareness raising activity in which students compared their academic introductions in English and in their first language, students started to note a wider range of problems other than sounds. Through the students' reflections their raised awareness becomes apparent. For example,

S10: *Although I try to pronounce the words clearly and imitate the intonation of native speakers, I found it not obviously shown on my video.*

S15: *I feel I use strong Chinese accent when I talk in English and I really want to change it.*

S9: *I speak very fast in English, which leads to a bad pronunciation.*

S11: *It seems I am not used to speak English. I am unfamiliar with the pronunciation of many words such as "improve" and "select", and I feel hard to tell the difference between [n] and [l]. I placed the stress on "detail" wrongly, and I did not use connections at all.*

Several weeks later, after learning about suprasegmental features and connected speech phenomena in class and working with the YoYoMa recording activity (described in Table 1 above), students' reflections continue to show expanded awareness and understanding of the

target speech features and how to improve them. For example, the following reflections illustrate students' thoughts on the process and purpose of the activity, as well as their conclusions on how the activity helped them note what to do to improve.

> S2: *This practice is helpful since I could get familiar to these thought groups, linkers, and stress. We imitate what the speaker says, and then record what we speak. By this method, we could compare our speech to the speaker and find what we should improve.*

> S14: *I have to say I like this practice. Doing this practice let me catch the breath points and intonation while listening to others. It's helpful to my self-training (...) Analyzing it before imitation makes me be able to follow the details of the talk. (...) I especially like the linking marked in the answer sheet. It helps me to find cues about what kinds of phrase I should link them together.*

Other students reflected about what they learned regarding the target features, what they were doing wrong with them, and what they could do to improve them.

> S12: *What I learn from this activity is stress. Before that I did not notice my problem on stress. I focused only on the fluency but not stress. (...) Now when I am doing listening practice or talking with natives, I pay much more attention to the stress...*

> S4: *Also, I found out that intonation can be changed even for the same phrases. Because the recording is still fast for me to took notes on the intonation, I made some "prediction" using my speaking habit and found out it is different from Ma's when I checked them again. Finally, I realized it is because he wants to emphasize a specific point in his lecture, so he makes that intonation and pause. Intonation is highly context-dependent.*

Students' discoveries and increased awareness of the target features and ways of improving them are also evident in their weekly pronunciation trackers. The following excerpts from students' weekly pronunciation trackers show what students learned regarding the target features after engaging in self-selected practice at home.

> S9: *I have learnt a lot form the record. I know that I should arrange the words in chunks. I also notice that I can use a long pause if I want to grasp other people's attention.*

> S11: *The division of thought groups is due to the meaning.*

> S6: *I found the intonation and rhythm very important for audience to understand after this practice.*

The above sample reflections indicate that the scaffolded activities, which included the explicit instruction of pronunciation strategies (focused on perception, production, and prediction skills), increased students' awareness and understanding of suprasegmental features and linking. These reflections and others also demonstrate that student-generated materials and practice activities for out-of-class practice resembled the activities, resources, and practice behaviors modeled and scaffolded by the teacher during the course (i.e., analyzing before reading aloud, recording, listening and imitating, self-assessing, etc.). For example, S9 noted that he chose to do the following during his self-practice:

> S9: *I read the news on the website and paid lots of attention to my intonation and linking. I listen carefully to the speakers' linking, stress and pause. Then imitate.*

It was therefore concluded that student-generated materials and practice behaviors suggested evidence of teacher scaffolding on the targeted features.

DISCUSSION

Before instruction, students seemed to know little about their pronunciation problems beyond specific sounds. It is possible that students think of the term pronunciation as equivalent to segmentals, which could explain why over half their responses indicated problems with sounds. It is also likely, however, that the students had limited knowledge and understanding of suprasegmental features or connected speech phenomena, or that they had never considered their own specific challenges with these targets. Lack of knowledge or lack of formal instruction on these features could also account for the limited descriptions initially. Regardless, from their initial responses, it seemed students had limited awareness of their pronunciation challenges in relation to suprasegmental features and/or linking. After students were asked to self-assess their production by listening to a recording of their speech and comparing it with their native language performance, they formulated richer descriptions of their pronunciation challenges, became more motivated to engage in pronunciation practice, and exhibited a greater understanding of what they needed to improve. This change in their motivations and perceptions provides validity to the role self-assessment plays in awareness raising and a rationale for the importance of integrating such practices into pronunciation training activities.

After receiving instruction on suprasegmentals and linking, and engaging in guided practice using a model speaker and script, students' reflections showed their increased knowledge and understanding about the target speech features. Their comments identified ways of improving the features and their recognition of their ability to produce them. In this activity, the instructor scaffolded learning by raising students' awareness of the target features (through explicit instruction, listening perception, and marking scripts), guiding their practice, giving them feedback, and asking them to reflect on the process and activity. Whether or not students' improvement in reduction, linking, stress placement, and intonation could be attributed to the instructor's scaffolding efforts could be debated, but the significant improvement and students' reflections provide evidence for teacher scaffolding being effective in producing a change in students' practice and production of the targeted speech features.

Interestingly, the instructor allotted comparably less class time to teach and practice vowel reduction than to teach and practice primary phrase stress, intonation, and linking. The little emphasis on vowel reduction (which was a conscious decision by the instructor) may explain why no comments were made regarding this specific feature in student reflections and self-assessments, and why students made the least improvement with this feature (+5.7%), achieving an overall accuracy of roughly over 50% at the end of the course. It is apparent that most students did not become aware that they needed to improve in reduction nor did they learn ways of doing so. This finding also seems to suggest the integral role of teacher scaffolding in establishing priorities for practice, modeling practice behaviors, and guiding students in the selection of features to target as well as in the selection of appropriate resources, activities, and pronunciation strategies. When the teacher did not scaffold learning for a particular feature (in this case, reduction), students' improvement was comparably less effective (+5.7%), and students' reports of practice with this feature outside of class were not found. On the contrary, improvements with primary phrase stress (+14.5%), intonation (+11.44%), and linking (+12.9%) were much larger and comments about work with these features outside of class were often present in students' self-reports. Because the features under study require somewhat different

pronunciation skills, this finding needs more validation from the field. However, it provides some evidence for teacher scaffolding playing a role in students' degree of pronunciation improvement and their choice of activities for self-improvement outside of class.

CONCLUSION

There are a multitude of factors involved in effective pronunciation training. This study provides evidence of one contributing factor: teacher scaffolding. Students' reflections and practice behaviors revealed that their improvement in the read-aloud diagnostics could be attributed at least in part to the role played by the teacher in scaffolding student learning; that is, in intentionally leading students' practice behaviors through carefully crafted activities that supported their learning and assisted their self-regulated efforts. The instruction was grounded in three pedagogical principles: prioritization of pronunciation practice based on student needs; empowerment through scaffolded instruction; and opportunities for students to monitor their performance and reflect on their outcomes (Sardegna & McGregor, 2012b). More research on teacher scaffolding is needed to better understand the instructors' role in facilitating effective pronunciation training.

ABOUT THE AUTHORS

Veronica Sardegna is an assistant professor in the Department of Curriculum and Instruction at the University of Texas at Austin. She is author of *The Practiced Business Writer: An ESL/EFL Handbook* (Briefings Media Group), board member of Tex-TESOL III, and Chair-Elect (2013-2014) for TESOL Speaking, Pronunciation, and Listening Interest Section (SPLIS). Her research focuses on learner autonomy, learning strategies, and second/foreign language learning and teaching. She is particularly interested in evaluating the effectiveness of using language learning strategies as well as instructional technology tools for autonomous language learning and teaching in the areas of English pronunciation, business writing, and oral communication. Contact her at vsardegna@utexas.edu.

Alison McGregor is an instructor for the International Office at the University of Texas at Austin. She teaches ITA and graduate level oral communication skills courses. In addition, she works as a communication consultant specializing in accent modification for the McCombs School of Business MBA program. Her research focuses on comprehensibility, language learning strategies, and factors that promote effective pronunciation training. Alison can be reached at mcgregor@utexas.edu.

REFERENCES

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge: Cambridge University Press.

Chamot, A. U. (2005). Language learning strategy instruction: Current issues and research. *Annual Review of Applied Linguistics, 25,* 112-130.

Chen, Y-M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research, 12*(2), 235-262.

Cohen, A. D. & Macaro, E. (2007). *Language learner strategies.* Oxford, UK: Oxford University Press.

Couper, G. (2011). What makes pronunciation teaching work? Testing for the effect of two variables: Socially Constructed Metalanguage and critical listening. *Language Awareness, 20*(3), 59-182.

Derwing, T. M., Munro, M. J, & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*(3), 393-410.

Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System, 30,* 155-166.

de Saint Lèger, D. (2008). Self-asessment of speaking skills and participation in a foreign language class. *Foreign Language Annals, 42*(1), 158-177.

Dickerson, W. B. (1994). Empowering students with predictive skills. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New views, new directions* (pp. 17-33). Alexandria, VA: TESOL Publications.

Dlaska, A, & Krekeler, C. (2008). Self-assessment of pronunciation. *System, 36,* 506-516.

Fitzgerald, J., & Graves, M. F. (2004). Reading supports for all. *Educational Leadership, 62*(4), 68–71.

He, L. (2011). Metacognition in EFL pronunciation learning among Chinese tertiary learners. *Applied Language Learning, 21*(1&2), 1-27.

Ingels, S. A. (2011). *The effects of self-monitoring strategy use on the pronunciation of learners of English.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign. Retrieved from http://hdl.handle.net/2142/26259

Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *The Modern Language Journal, 90*(iii), 320-337

Morley, J. (1994). Multidimensional curriculum design for speech-pronunciation instruction. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New views, new directions* (pp. 64-91). Alexandria, VA: TESOL Publications.

Olson, C. B. & Land, R. (2007). A cognitive strategies approach to reading and writing instruction for English language learners in secondary school. *Research in the Teaching of English, 41*(3), 269-303.

Osburne, A. G. (2003). Pronunciation strategies of advanced ESOL learners. *IRAL, 41*, 131-143.

Sardegna, V. G. (2009). *Improving English stress through pronunciation learning strategies.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign (UMI No. 3363085).

Sardegna, V. G. (2011). Pronunciation learning strategies that improve ESL learners' linking. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 105-121), Ames, IA: Iowa State University.

Sardegna, V. G. (2012). Learner differences in strategy use, self-efficacy beliefs, and pronunciation improvement. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2011. (pp. 39-53), Ames, IA: Iowa State University.

Sardegna, V. G., & McGregor, A. (March 2012a). *Enhanced self-assessments for pronunciation improvement.* Paper presented at the 46th Annual TESOL Convention and Exhibit, Philadelphia, PA.

Sardegna, V. G., & McGregor, A. (2012b). Principles for teaching pronunciation to international teaching assistants. *TESOL SPLIS Newsletter*, *8*(1).

Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing, 15*(1), 1-20.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching how to listen makes a difference: An empirical study. *Language Learning, 60*(2), 470-497.

Walqui, A. (2006). Scaffolding instruction for English learners: A conceptual framework. *The International Journal of Bilingual Education and Bilingualism, 9*(2), 159-180.

PRONUNCIATION TEACHING AND LEARNING: EFFECTS OF EXPLICIT PHONETIC INSTRUCTION IN THE L2 CLASSROOM

Joshua Gordon, Indiana University

Isabelle Darcy, Indiana University

Doreen Ewert, University of San Francisco

This classroom-based study investigated how explicit instruction influences the acquisition of L2 phonological features, and how their production contributes to comprehensible speech in L2 learners. Three groups of ESL learners (n=30) received pronunciation instruction during three weeks (25 minutes/day, 3 days/week) using the same teaching sequence within a communicative methodology (Celce-Murcia, Brinton, & Goodwin, 2010; Hinkel, 2006). Groups differed in the type of explicit instruction received: two experimental groups received explicit instruction either on suprasegmental or segmental features. A combination of these features was presented orally to the third group without explicit instruction. Pretest and posttest recordings indicate that only the explicit group trained on suprasegmentals improved its comprehensibility scores significantly from pretest to posttest. An analysis of classroom-treatment recordings demonstrates that explicit phonetic instruction that makes learners notice L2 features (i.e., explicit presentation of contents, guided analysis and practice, and corrective feedback) can be beneficial for L2 learners in the development of comprehensible speech. The results also corroborate a major role of prosody in enhanced comprehensibility (e.g., Derwing, Munro, & Wiebe, 1998; Munro, 1995) and a call for more explicit phonetic instruction within a communicative methodology (e.g., Celce-Murcia et al., 2010).

Different pronunciation studies have demonstrated the benefits of explicit phonetic instruction in pronunciation learning (e.g., Lord, 2005; Saito & Lyster, 2012a, 2012b). However, the effects of instruction targeting segmentals or suprasegmentals in comprehensibility (e.g., Derwing, Munro, & Wiebe, 1998) or the application of laboratory training studies to classroom instruction (e.g., Pennington & Ellis, 2000) has not been investigated in large intensive language programs where the implementation of pronunciation instruction is sometimes limited. This classroom-based study investigated how explicit instruction influences the acquisition of second language (L2) segmentals or suprasegmentals, and how their production contributes to comprehensible speech in L2 learners in a large intensive language program. The study was implemented in three intact English as a second language (ESL) classes where learners received explicit instruction in either segmentals or suprasegmentals, or no explicit instruction. This treatment implementation took place during three weeks (three days, 25 minutes each day), and the results suggest a major role of prosody in enhancing comprehensibility in learners even when instruction is limited to small periods of time in class. Additionally, a qualitative analysis of the classroom-treatment implementation demonstrated that explicit phonetic instruction that guides learners to notice L2 features—and the ambiguities that could arise in production as a result of errors—could be beneficial in the development of comprehensible speech. We first present some background information that motivated our study as well as our methodology followed by our discussion and conclusions for pronunciation teaching.

LITERATURE REVIEW

Nonnative pronunciation is perceived in the production of both segmentals and suprasegmentals in L2 speech, it contributes to the perception of foreign accent, and it may lower intelligibility or comprehensibility in speech (Kang, Rubin, & Pickering, 2010; Munro & Derwing, 2008; Trofimovich & Baker, 2006). Additionally, nonnative production of suprasegmentals appears to be more detrimental than segmental errors in L2 comprehensibility and intelligibility perception (see Field, 2005; Kang, et al., 2010). To help L2 learners with these problems, training studies have proven to be beneficial in speech perception/production. For instance, high variability training studies have shown improvement in learners in both perception and production of segmentals and suprasegmentals (see Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1997; Wang, Spence, Jongman & Sereno, 1999; Wang, Jongman & Sereno, 2003). Additionally, Pennington and Ellis (2000) demonstrated that directing learners' attention to and raising their awareness of prosodic features of the L2 during training improved their interpretation of sentence meaning. These results are significant because they call for a stronger role of phonetic explicitness in L2 pronunciation instruction in classroom settings.

In pronunciation teaching, explicit phonetic instruction has demonstrated positive benefits (Lord, 2005) and instruction on suprasegmentals appears to yield better improvements in comprehensibility as opposed to instruction on segmentals only (see Derwing, et al., 1998). Moreover, researchers have also pointed out that L2 learners might not necessarily put into practice in spontaneous speech what they learn under controlled tasks (Celce-Murcia, Brinton, & Goodwin, 2010; Bowen, 1972). This is why a communicative component in pronunciation instruction has been advocated to develop fluent and comprehensible L2 speech (see Celce-Murcia, et al., 2010; Hinkel, 2006).

In spite of the research evidence, one of the real challenges in instruction is to bridge the gap between theory and practice given the disconnection between research in L2 phonology and the real practices in the classroom (see Derwing & Munro, 2005; Levis, 1999). It is also the case that in pronunciation teaching, the communicative framework has often been perceived as incompatible with explicit pronunciation instruction (Darcy, Ewert, & Lidster, 2012; Derwing & Foote, 2011) – even though explicit instruction is necessary to develop accuracy, which is a key factor in communicative competence.[1] Additionally, only a few studies have tried to apply the findings of laboratory phonology research to L2-classroom practices (see Derwing et al., 1998). Therefore, more research that integrates the findings from laboratory studies into real, time-constrained L2 pronunciation instruction—within a communicative methodology—is necessary, given the potential benefits it could bring to learners in the development of comprehensible L2 speech.

THE CURRENT STUDY

Given the importance of comprehensible speech in pronunciation instruction at present (Levis, 2005, 2006),[2] we investigated the development of comprehensibility in L2 learners through

---

[1] Accuracy, or "grammatical competence" in the syntactic and phonological levels is one of the cornerstones of communicative competence. See Savignon (2000) for a review of communicative competence in language teaching.

[2] *Comprehensibility* refers to a listener's estimation of difficulty in understanding an utterance produced by an L2 speaker, and it is usually measured using some form of scale (Munro, Derwing, & Morton, 2006). Levis (2005, 2006) explains that both *intelligibility* and *comprehensibility* are related terms. Intelligibility, in the broad sense given by Levis (2006), refers to the listeners' ability to understand speech and "is not usually distinguished from

explicit phonetic instruction within a communicative methodology. The study was guided by the following research questions:

1. Does short-term, explicit pronunciation instruction increase comprehensibility more than nonexplicit instruction?

2. Does instruction in suprasegmental features increase comprehensibility more than instruction in segmental features?

The details of instruction and the measurements of comprehensibility in speech through a pretest and a posttest are explained in the next sections.

Participants

Three intact intermediate speaking classes (Level Six out of seven institutional levels) in an intensive English program at a large American university participated in the study (see Table 1). Thirty students distributed among these three classes were the potential participants in the study. Native speakers of American English also participated in the study in three different ways. First, two speakers (male and female) prepared the pre and posttest sentences used as prompts. Second, another group of 10 speakers recorded the same sentences as the learners, in the same conditions, for comparison purposes. These speakers recorded the sentences only once. Finally, an additional group of 12 speakers participated in the comprehensibility rating of the learners' productions after the posttest. These raters were graduate students in linguistics or language teaching.

Table 1
L2 Learners and L1 English Participants

Participants	Group	Condition	*n*	TOEFL Score
Speakers	1. L2 Learners	Suprasegmental	12 (4)	499.41
	2. L2 Learners	Segmental	8 (4)	514.22
	3. L2 Learners	Nonexplicit	10 (4)	484.85
	4. L1 Speakers	Baseline	10 (4)	
Listeners	L1 Speakers	Comprehensibility rating	12	

Note: Numbers in the fourth column represent the original number of students in each class as well as the actual number of students (in parentheses) who were included in the final analyses.

Instruction

This study followed a pretest-posttest experimental design. The two treatment groups (i.e., suprasegmental and segmental groups hereafter) received treatment during three weeks, three days per week, and 25 minutes each day (total: 225 minutes of instruction). To implement a communicative methodology, each lesson followed a presentation-practice-production sequence

---

closely related terms such as comprehensibility" (p. 252). However, for the purpose of this study, we will use the term *comprehensibility* as the main goal to attain in pronunciation instruction.

(Chamot, Barnhardt, El-Dinary, & Robbins, 1990). This was used so that teachers could introduce phonetic content explicitly, guide the students in controlled tasks, and then provide room for a communicative activity where learners could produce and put into practice each lesson's content.

Each experimental group received explicit phonetic instruction in either specific suprasegmentals (i.e., stress, rhythm, linking, reductions) or segmentals (i.e., vowels /i, ɪ, æ, ɛ/). We targeted these specific segmentals and suprasegmentals because pronunciation materials have pointed out the difficulty they pose for different L1 groups when learning English (see Avery & Ehrlich, 1992). The third group (i.e., nonexplicit group) did not receive explicit pronunciation instruction but engaged in the same practice and production activities of the two treatment groups. Table 2 presents the details of the treatment used.

Table 2
Classroom Treatment Description

Stages and Techniques	Class 1 (*n*=12) Suprasegmentals	Class 2 (*n*=8) Segmentals	Class 3 (*n*=10) Nonexplicit
Presentation Visual aids Oral introduction of topic	Explicit instruction and analysis	Explicit instruction and analysis	Nonexplicit instruction, pronunciation practice announced
Practice Bottom-up skills, analysis, recognition & discrimination minimal pair drills, reading short passages	Rhythm, stress, reductions, linking.	Vowels /i, ɪ, æ, ɛ/ and articulation, vowel contrasts, minimal pairs	Classroom drills on words, phrases, and sentences; combination of the same materials as the other two groups
Production Top-down skills, fluency activities	Communicative tasks: Pair discussion, group discussion, role-plays, information-gap activities		

As seen in Table 2, the experimental groups received explicit phonetic instruction and analysis of content in the topic introduction stage. In contrast, the nonexplicit group only received an announcement that pronunciation would be practiced. For the practice stage, the experimental groups carried out different tasks, such as minimal-pair recognition and discrimination, or analysis of stress and rhythm in short passages and sentences. The nonexplicit group listened to and repeated words, phrases, and sentences during this stage.[3] Finally, the three groups had similar production tasks such as pair and group discussions, role-plays, or information-gap activities. To measure the effects of instruction in the development of comprehensibility, we collected pre- and posttreatment speech samples from the L2 learners who participated in the study, which we later presented to a group of L1 English speakers in a comprehensibility-rating

---

[3] These words, phrases, and sentences were taken from the materials used in the two experimental groups.

task.

All the treatment sessions were audio recorded, and one of the researchers sat in the three classes as a nonparticipant observer. The audio recordings were later broadly transcribed to analyze what happened in each class to verify implementation of the experimental design (see van Lier, 1988).

Pretest and Posttest

All the participants from the three classes were audio-recorded individually before and after treatment. The recordings took place in a sound-isolated booth in a psycholinguistics laboratory. The participants recorded sentences through a delayed-sentence repetition task using prompts previously recorded by two L1-speakers (male & female) of American English (see Trofimovich & Baker, 2006). There were 24 sentences for the pretest and 48 sentences for the posttest (i.e., the same 24 sentences from the pretest plus 24 new sentences for generalization purposes). These sentences were carefully designed to have at least a word with one of the four vowel sounds studied by the segmental group (i.e., / i, ɪ, æ, ɛ /), or function words (e.g., articles, prepositions) that could trigger vowel reduction—a key aspect in stress-timing and rhythm, which were studied by the suprasegmental group.

Given the classroom-based nature of this study, attrition was very high because of two factors. First, a high number of participants from the three L2-groups missed classes on a regular basis or enrolled later in the course, and second, the delayed-sentence repetition task was challenging for the majority of L2-learners and many sentences were simply not accurately produced. Because of these problems, for our final analysis we selected only speech samples from four participants from each of the three classes who completed the treatment sessions. Therefore, in the final analysis we only included (a) sentences that were produced by those L2 learners who completed the treatment in its entirety, and (b) sentences that were produced correctly in both tests. In order to have comparable numbers across groups, we only used sentences from four L1-English speakers (4 L2-participants x 3 groups = 12 participants + 4 L1-English speakers = 16 speakers). This gave us a total of eight pretest sentences and sixteen posttest sentences per participant to use in the comprehensibility-rating task (8 pretest + 16 posttest = 24 sentences per participants, totaling 384 sentences). These sentences were presented to a group of twelve L1-English listeners to be rated for comprehensibility. These raters were graduate students in linguistics or language teaching. In the rating task, six raters listened to half of the total amount of sentences and the other six raters listened to the other half to avoid fatigue. The raters listened to the sentences in a computer lab through high-quality headphones. Each sentence was rated using a 9-point Likert scale (1=*extremely easy to understand*, 9=*impossible to understand*) similar to the one used in other studies and which has proved to yield highly reliable ratings (see Derwing & Munro, 1997; Munro & Derwing, 1995). The results of this comprehensibility task are presented below together with a qualitative analysis of the implementation of instruction in class.

RESULTS

Quantitative

The inter-rater reliability coefficients (Cronbach's alpha) computed across all ratings given for each list were high (.92 and .92), which indicated very strong agreement (LeBreton & Senter, 2008). Learners obtained equal performance for repeated and new sentences at the posttest (no effect of *sentence type* [new vs. repeated] and no interaction with *group*, both $p > .05$), allowing us to collapse new and repeated sentences at posttest for the comparison with pretest ratings.

Table 3 shows the average rating obtained for each group at each time. Only the suprasegmental group significantly improved from pretest to posttest in comprehensibility, their mean rating being closer to the L1 speakers'.

Table 3
Average ratings for pretest and posttest for each group

Group	Pretest	(*SE*)	Posttest	(*SE*)	Significant difference?
Suprasegmental	4.2	.39	3.6	.38	*
Segmental	4.5	.39	4.8	.38	*
Nonexplicit	4.4	.38	4.2	.38	>.05
L1 Speakers	1.1	.38	1.1	.38	>.05

Note: *SE* is the standard error of the mean.

A global mixed effects model declaring the factors *test* (pretest and posttest), and *group* (suprasegmental, segmental, nonexplicit, L1-English speakers) was conducted with speakers, token, and raters as random effects. There was no main effect of *test* ($F_{1,1371} = 2.3$, $p > .1$), but a large effect of *group*, ($F_{3, 12} = 39.0$, $p < .001$), and a significant interaction of *group* and *test* ($F_{3, 2032} = 7.5$, $p < .001$) suggesting that the groups received different ratings at each time, and that this difference was modulated by the kind of treatment received. The interaction remained significant when restricting the analysis to L2 learners only ($F_{2, 1649} = 9.7$, $p < .001$), indicating that the performance at each test varied as a function of the treatment received.

Qualitative

It is important to stress that we provided the materials and lesson plans to the three collaborating teachers in this project, but their implementation in class was left to their own teaching preferences. We codified themes and categories using a comparative method (Glasser & Strauss, 1967; Richards, 2003). The analysis revealed important findings in terms of explicit versus nonexplicit instruction, such as clarifications of ambiguities caused by mispronunciations, and reinforcements of concepts through feedback and explanations.

One of the clearest differences between the experimental and nonexplicit groups confirmed that the level of explicit phonetic information given by the teachers to signal possible communication problems was higher in both experimental groups—as intended in our research design. For example, both teachers in the experimental groups introduced topics emphasizing the ambiguities that mispronunciations could create in meaning. Two passages presented in Appendixes 1a and 1b show the teachers introducing word stress and the contrast /i-ɪ/ in class. In contrast, the nonexplicit group did not receive this type of instruction, and the teacher only gave the students an announcement of pronunciation instruction followed by asking the students to listen and repeat, as shown in Appendix 1c.

It is also important to stress that teachers in the experimental groups were consistent in making learners aware of how mispronunciation could create communication problems not only when introducing content, but also in controlled and communicative tasks. This was not the case in the nonexplicit group where learners' attention was not directed to these issues. Appendix 2a shows the teacher in the suprasegmental group instructing the students to carry out a conversation in

groups of three, but taking turns monitoring each other and paying special attention to prosodic aspects analyzed before. Similarly, Appendix 2b shows the teacher in the segmental group pointing out communication problems when mispronouncing /i-ɪ/ in a very common minimal pair. In contrast, even in communicative activities in the nonexplicit group, the teacher directed the students to carry out the task (like a conversation in groups), but there was no information about phonetic elements or feedback on the learners' pronunciation—as shown in Appendix 2c.

The teachers in the two experimental groups also reinforced content while the students worked in pairs or groups. This was done by providing feedback, making comprehension checks, or assisting the students individually when there were very specific problems with either segmentals or suprasegmentals. On the contrary, the students in the nonexplicit group did not receive this type of reinforcement. Instead, most of their attention during instruction was focused on meaning and not on any analysis of phonetic phenomena. Although the students in both experimental groups also had communicative activities where meaning was important—especially in regard to the differences in meaning that mispronunciations could create—the teachers also raised the students' awareness that these ambiguities in meaning can be the result of problems with the pronunciation of the vowels or suprasegmentals studied. These differences are seen in Appendices 3a and 3b, where the teacher in the suprasegmental group helps a student with the stress in a sentence while working in pairs (3a), or the teacher in the nonexplicit group asks the students to share comments on issues discussed previously—in terms of meaning but not on phonetic information (3b).

DISCUSSION AND CONCLUSIONS

The quantitative analysis of the rating task and the qualitative analysis of the classroom implementation provide compelling evidence for the role of form-focused instruction for pronunciation in the communicative classroom. First, in the suprasegmental group, the explicit phonetic information regarding possible miscommunication issues that come up as a result of suprasegmental problems and directed feedback on production resulted in improved comprehensibility ratings. This was not the case in the nonexplicit group where the students only repeated words and phrases or had activities that emphasized fluency. This supports research which has shown that a lack of focus on form can develop fluency in learners but not necessarily accuracy in their production (see Lyster & Ranta, 1997) as well as research that indicates the positive effects of feedback in pronunciation instruction (e.g., Saito & Lyster, 2012a & b). Together, the evidence that directing learners' attention to prosodic forms enhances interpretation of sentence meaning (see Pennington & Ellis, 2000), and the evidence from this study that directing learners' attention to linguistic features that often obscure meaning improves production provide a strong rationale for the inclusion of components of explicit pronunciation instruction in the communicative classroom.

The differences in the comprehensibility ratings of the two experimental groups, however, raise questions of how best to introduce different types of pronunciation content. In general terms, the suprasegmental group appears to have had a more global, complex, and explicit analysis of the suprasegmentals content than the segmental group did with its content. The suprasegmental group analyzed rhythm, stress, linking, and reductions in words, phrases, sentences, and more complex discourse units such as short passages. On the other hand, the segmental group focused its analysis on only four vowel sounds, and this was primarily done in the meaning differences among lexical items. As demonstrated by previous research, this probably did not fully engage the attention capacity of learners in the segmental group by only focusing on the pronunciation

of those vowel sounds and not other global aspects that also affect speech perception (see Derwing et al., 1998; Schmidt, 2001). These differences may account for the ratings, but more controlled research will be necessary in the future in order to determine if this is an adequate analysis.

In conclusion, the results of this study point out that explicit phonetic instruction benefits L2 learners overall, which confirms previous results (e.g., Derwing et al., 1998; Levis, 1999; Lord, 2005). These results demonstrate that even adding only a relatively time-limited explicit pronunciation component in a primarily communicative classroom can lead to beneficial results in production for learners. In addition, the qualitative analysis demonstrates that complex classroom interactions of instructional focus and teacher implementation can make a difference in learner development, as demonstrated in the differences in explicitness between the experimental and nonexplicit groups in this study. Finally, although segmental instruction is indeed necessary and important, the difference in outcomes between the two treatment groups suggests that explicit suprasegmental instruction may give learners a faster improvement in comprehensibility rather than only focusing on a few vowel sounds.

ABOUT THE AUTHORS

Joshua Gordon is a doctoral candidate in the Department of Second Language Studies at Indiana University. He has taught ESL, EFL, and Spanish as a foreign language. His research interests include second language phonology and second language pedagogy, particularly the interface between both areas in the teaching and learning of pronunciation skills.

Affiliation: Indiana University
Address: 1021 E 3rd St, Memorial Hall 315. Indiana University, Bloomington, IN 47405.
Email: jgordonz@indiana.edu

Isabelle Darcy is Associate Professor in the Department of Second Language Studies at Indiana University, where she teaches graduate courses in methodologies for teaching pronunciation. She obtained a Ph.D. in Linguistics and Cognitive Science from the Ecole des Hautes Etudes en Sciences Sociales in Paris (France) and from the Johannes Gutenberg University in Mainz (Germany), for which she was given an Outstanding Dissertation award. Her research focus is the acquisition of second language phonology, as well as native and nonnative speech perception and word recognition.

Affiliation: Indiana University
Address: 1021 E 3rd St, Memorial Hall 301. Indiana University, Bloomington, IN 47405.
Email: idarcy@indiana.edu

Doreen Ewert is Associate Professor and Director of the Academic English for Multilingual Students Program in the Department of Rhetoric and Language at the University of San Francisco. She has taught ESL/EFL learners, trained SL/FL teachers, and directed ESL programs for pre-matriculated and matriculated adult learners for several decades. She obtained a Ph.D. in Linguistics at Indiana University. Her work includes language curricular design and implementation, pre- and in-service language teacher development workshops and adult SL literacy development. Her research focus is on SL/FL reading and writing development, curriculum implementation, and teacher development.

Affiliation: University of San Francisco

Address: 130 Fulton St., Kalmanovitz Hall 205. University of San Francisco, San Francisco, CA 94117.
Email: dewert@usfca.edu

REFERENCES

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation*. New York: Oxford University Press.

Bowen, J. D. (1972). Contextualizing pronunciation practice in the ESOL classroom. *TESOL Quarterly, 6,* 83-94.

Bradlow, A., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics, 61,* 977-985.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. (2nd ed). New York: Cambridge University Press.

Chamot, A., Barnhard, S., El-Dinary, P., & Robbins, J. (1999). *The learning strategies handbook*. London: Longman.

Darcy, I., Ewert, D. & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox." In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, (pp. 93-108). Ames, IA: Iowa State University.

Derwing, T. & Foote, J. (2011). 2010 National survey of pronunciation teaching: Deja vu. Paper presented at the Conference of the *Annual Association for Applied Linguistics*, Chicago, IL, March 26, 2011.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19,* 1-16.

Derwing, T. M., Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39,* 379-397.

Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48,* 393-410.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39,* 399-423.

Glasser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine.

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly, 40,* 109-131.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *Modern Language Journal, 94,* 554-566.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*, 815-852.

Levis, J. M. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly, 33*, 37-63.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, *39*, 369-377.

Levis, J. M. (2006). Pronunciation and assessment of spoken language. In R. Hughes (Ed.) *Spoken English, TESOL and applied linguistics* (pp. 245-270). New York: Palgrave Macmillan.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania, 88,* 557-567.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition, 19,* 37-66.

Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition, 17,* 17-34.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45,* 73-97.

Munro, M. J. & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning, 58,* 479-502.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28,* 111-131.

Pennington, M. C. & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *Modern Language Journal, 84,* 372-389.

Richards, K. (2003). *Qualitative inquiry in TESOL.* New York: Palmgrave Macmillan.

Saito, K., & Lyster, R. (2012a). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning, 62*, 595-633.

Saito, K., & Lyster, R. (2012b). Investigating pedagogical potential of recasts for L2 vowel acquisition. *TESOL Quarterly, 46*, 387-398.

Savignon, S. J. (2000). Communicative language teaching for the twenty-first century. In M. Celce-Murcia (Ed) *Teaching English as a second or foreign language.* Boston, MA: Heinle & Heinle.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.). *Cognition and second language instruction* (pp. 3-32). New York: Cambridge University Press.

Trofimovich, P. & Baker, W. 2006. Learning second language suprasegmentals: Effects of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition, 28,* 1-30.

van Lier, L. (1988). *The classroom and the language learner.* New York: Longman.

Wang, W., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluations of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America, 113,* 1033-1043.

Wang, W., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America, 106,* 3649-3658.

Appendix 1

a. Suprasegmental Group:

T: But, [teacher reading from the slide] syllables and words in English contain stress. This means that some syllables in a word, or some words in phrases and sentences are produced longer, louder, and higher. That's what I'm used to listening to, and I would need to hear that stress in order to understand the important information, ok? So, we're going to be learning about that, and… stress can make a difference in the meaning of words and phrases, so for example, if I say this, what is this??

Ss: record…
T: REcord, RE-cord [emphasizing the stress], so the stress is on the first syllable. RE-cord. I'm talking about this [pointing to the picture of a record in the slide]. If… I go over here and I say, re-CORD! re-CORD! then it's different. It means somebody talking into a microphone and recording their voice. So this one is actually a noun, and this one is the verb. Even though they're spelled exactly alike, the stress is different. Ok? And you'll learn as we go along that often with nouns, the stress is on the first syllable, and that same word if it's a verb, the stress shifts to the second one.

b. Segmental Group:

T: So the first one, um… (3s) *peel* like when you have an orange, the hard part on the outside that you don't eat, this is *peel* [teacher shows students the slide with pictures of an orange *peel* and a *pill*. She also emphasizes the vowel pronunciation]
Ss: peel [students repeat]
T: It makes you, feel, feel your, feel your lips going this way, peeeeeel! [teacher emphasizes the pronunciation]
Ss; peel, peel [students repeat]
T: Ok, this like medicine… one… little thing, is a… "pill" [teacher emphasizes the relaxed quality of the vowel sound]
Ss: pill [students repeat]
T: so your cheek should be relaxed, "pill"….
Ss: pill
T: so this one is pEEl [pil]
Ss: peel
T: Pill [pɪl]
Ss: pill

c. Nonexplicit Group:

T: All right, um… so most of this is kind of listen, listen and repeat, and then we'll have, um… some kind of group activity at the end, ok? Questions?
Ss: no
(…)
T: All right, ready?
Ss: [students mumble]
T: Ok, I'm going to read these words, ok? and then you just, um… repeat… right? so yeah,… Father…
Ss: Father…
(…)
T: Advantage…
Ss: Advantage
T: Actually I don't pronounce the "t," "advantage…"
Ss: Advantage…
T: [laughs] Playing…

Appendix 2

a. Suprasegmental Group:

T: So student A will talk, but student B will also try to ask questions and draw out information from student A. Ok, now student C, you're very important…
Ss: Ohhhhhh….
T: You're going to listen, to what student A says, and make sure that there are a few reductions. So, of course you will hear stress, but… try to think… is this person… [claps and exaggerates her voice] stressing every word, because we don't want that. Can you hear spaces between the stresses? I have a MEmory, I have a MEmory [teacher claps], there's got to be real spaces. Ok, if possible, also listen for stress… and the use of rhythm groups… do you remember rhythm groups? Those are the pauses… 'I have a memory… it was when… I was two years old…' so, is there some pauses? Can you hear spaces, between the stressed words?? Ok, then you'll change roles, and everyone will have an opportunity to tell a story, and to monitor. Ok, so, do you understand? (3.s)
S?: Mmm-hmm
T: Who's person A??
Ss: (…)
T: you talk first, you help them talk, you listen [talking to the different members of a group], ok? you just talk for about a minute, and then we will hear your (…), just concentrate on (…), then you will talk… all right? And you will listen! Got it? [students work for the rest of the class on this. The teacher walks around helping groups when necessary.]

b. Segmental Group:

T: live
Ss: live
T: leave
Ss: leave
T: I gotta tell you, this is probably the most important pair [pointing at live, leave], many many many many times do you remember like in week 1, when you had to interview your classmates, and people would say 'where-do-you-leave??'… and the other student is like… "I'm not leaving!"
Ss: [laughter]
T: you know? And there's also confusion and they say, you know, "Korea!" and they're like, like… 'cause the question, they don't know "where did leave?" or "where do you live?", so now you "live" [exaggerating] everything is relaxed, do it, now you 'live' in Bloomington. For Spring break maybe you… left…. well that's past tense anyway… but Summer vacation, after Spring II, before Summer I, maybe you'll "leave!" [exaggerating] to go home, but now you "live" [exaggerating], relaxed…

c. Nonexplict Group:

T: All right, we have a little exercise, um… it is in pairs… (4.s) [puts the new slide on the projector], so in pairs, tell a classmate about your experience learning English in the United States. Make sure you provide enough details, so, um… you can explain the following issues: 'what things have been easy and which ones have been difficult… in learning English… so grammar, pronunciation, reading and writing… vocabulary.' 'Have you had a funny experience with the language? for example, pronouncing something inappropriately…' you have??
Ss: [laughter]
T: [laughs] yeah! 'What is something you always forget in English… you can share with your classmate… um… so let's work in… pairs…

[students work in pairs for about 8 minutes while the teacher walks around listening to the students' stories and helping them with vocabulary]
T: Ok… [students keep talking in groups and the teacher helps a girl with some vocabulary] ok, um… if your partner had a funny experience… learning English… you can share it the rest…
S?: Noooo!!!! [laughs]
All: [laughter]
T: Well if you're willing…

Appendix 3

a. Suprasegmental Group:

T: …oh which one??, *She's a cool teacher*, *SHE's a cool teacher! She's a COOL teacher?? She IS a cool teacher"* Notice the difference in meaning [not able to understand] … that means I'm stressing this, *She IS a cool teacher*, and it can be negative too, *She ISN'T a good teacher*. That means she is not a cool teacher. But, I'm trying to make it sound very positive, *She IS a good teacher, She IS a good teacher*. If this were here a contraction, it would be '*She's*' *She's a cool teacher*, *She's a cool teacher*. But I'm trying to make everybody understand that this is positive…

b. Nonexplicit Group:

T: Ok, so first we're just going to listen and repeat, ok? "*Good afternoon*"
Ss: "*Good afternoon*"
T: "*Pill*"
Ss: "*Pill*"
T: "*Rock and roll*"
Ss: "*Rock and roll*"
T: "*It is a life or death matter*"
Ss: "*It's a life or death matter*"
T" "*It is as cold as ice*"
Ss: "*It is as cold as ice*"
T: Good
T: "*She's one in a million*"
Ss: "*She's one in a million*"
T: What does that mean, uh?
S: Oh, uh she's very good.
T: Yes, she's special, she's unique. "*Do you have his address?"*
Ss: "*Do you have his address?*"

INCORPORATING PRONUNCIATION IN THE FIRST-YEAR SPANISH CLASSROOM: AN EARLY INTERVENTION

Ann Aly Bailey, University of California, Los Angeles

Anel Brandl, The Florida State University

There is currently no consensus about the role of pronunciation instruction in beginning L2 classrooms. Trofimovich and Gatbonton (2006), and Hurtado and Estrada (2010) found that focusing on form and meaning resulted in greater improvement in the pronunciation of intermediate than beginning learners, and Hurtado and Estrada (2010), and Arteaga (2000) suggest that perceptual training could be beneficial for beginners.

This study investigates the effects of two types of instruction (Technical and Non-technical) on the perceptual abilities of beginning L2 Spanish learners. The Technical Instruction (TI) group received information on target sounds via formal phonetic explanations. The Non-technical Instruction (NI) group was aurally exposed to the target sounds and made aware of their orthographical contexts without formal phonetic explanations. Both groups completed perception activities with the target sounds. A control group of the same level received neither instruction nor perception activities. Assessments were online and consisted of a pretest, posttest and delayed posttest that measured perceptual accuracy of the target sounds.

In general, the results did not reveal a significant improvement for the groups in their perceptual development by the end of the experiment. However, there was a significant effect for instruction at the immediate posttest, in which the NI group outperformed the TI and Control groups. This effect was neutralized in the delayed posttest, which revealed no significant differences between groups. A further analysis by sound revealed gains on the vowels, voiceless stops and the grapheme 'v', which provides support for the claims of the Perceptual Assimilation Model-L2 (Best and Tyler, 2007).

As the field of second language acquisition (SLA) continues to grow and contribute important insights to the acquisition process, pedagogical approaches are being developed and modified to accommodate these findings. Researchers such as Van Patten (2004, 2007) and Swain (1995, 1998) suggest the importance of communicative and meaningful learning contexts, with Van Patten focusing on the type of input provided to learners and Swain focusing on learner output as a part of the acquisition process. Although many university-level foreign language programs acknowledge the findings of SLA and incorporate a communicative approach into their classrooms, this pedagogical method has focused more on the processing and acquisition of syntax and morphology. The aspects relating to pronunciation or the phonological system of the second language (L2) are rarely addressed in beginning-level communicative classrooms, although a certain level of accurate pronunciation is needed to achieve communicative goals. This linguistic deficit is acknowledged and discussed in several studies which attempt to examine the effectiveness of pronunciation instruction in the classroom, several of which are reviewed in the next section.

Previous empirical studies on pronunciation instruction

Various studies have incorporated phonetics instruction in L2 classrooms with positive results for both production and perception assessments, but the majority of them take place in intermediate-level (third and fourth year university level) courses. The techniques and results of several of these studies will be discussed below as well as possible modifications to make these procedures appropriate for a beginning (first year) L2 classroom.

Elliot (1997) examined the effectiveness of formal phonetic instruction for intermediate-level Spanish learners at an American university over a semester. Assessments included a pretest and posttest which included various production tasks. Participants' production was judged impressionistically by one native and two non-native speakers. Results from Elliot (1997) showed significant improvement for the experimental groups on the word reading and word and sentence repetition tasks, and improvement approached significance in the experimental participants' spontaneous speech, whereas there were no significant gains for the control group. In reference to improvement on groups of target sounds, the experimental group had significant gains on liquids and stops, and marginal gains on vowels.

Lord (2005) also used a formal Spanish phonetics course as the experimental setting. A pretest and posttest design required participants to read a paragraph from which the target forms were taken and acoustically analyzed in isolation. The treatment included instruction on technical linguistic terms, comparisons of the participants' spectrograms with those of native speakers, and oral practice. The results of the production tasks showed that L2 participants produced native-like VOTs for /p, t, k/ on the posttest (although there was not a significant gain from pre to posttest), and had significant gains from pretest to posttest on /r/, diphthongs, and the spirants [β, ð, ɣ].

Ausín and Sutton (2010) examined the progress of third year learners in a formal Spanish phonetics course using perception assessments. The treatment took place over a semester and used a pretest and immediate posttest design with two sections of a phonetics course. Typical class activities included explicit descriptions of target forms, transcriptions, and oral exercises. The assessment is a version of a grammaticality judgment task that was adapted to test the participants' perception based on its pronunciation, not its semantic or syntactic content. To complete the task, participants saw Spanish words on a computer screen and then heard their pronunciations twice, after which participants were asked to rate the pronunciation of the words they heard on a scale of 1 ("very bad") to 5 ("native-like"). Significant differences were found in judgments of the English-like pronunciations from pretest to posttest and for the graphemes 'z' and 'h', the spirants [β, ð, ɣ], word-final /l/, and the fricatives /h, v, z/.

Trofimovich and Gatbonton (2006) is one of the few studies on pronunciation instruction that included a beginning level group as a comparison to an intermediate level group of L2 learners. This study tested whether repetition and Focus on Form (FonF) and Focus on Meaning (FonM) tasks with pronunciation would show an improvement in either proficiency group. For the repetition task, the response time and judgments of the participants' spoken words revealed that they processed and pronounced repeated words faster and with more accuracy, indicating sensitivity to repetition.

This experiment used a FonF and a FonM task. To bring the participants' attention to either form or meaning, they responded to a different list of words by rating from 1-10 either the "degree of pleasantness" (p. 525) or spoken clarity of the word. Participants then repeated the word. Results

showed that the FonF task did not affect the participants' sensitivity to repetition, whereas the FonM task only affected the sensitivity of the lower-proficiency learners. Trofimovich and Gatbonton concluded that the lower-proficiency participants' results on the FonM task could be due to a mismatch between what they were attending to and producing.

The above studies all suggest a positive effect for pronunciation instruction at the intermediate level, although not on all forms or tasks elicited. Trofimovich and Gatbonton (2006) also had positive results with beginning learners, as long as they were not attending to both form and meaning. However, there were limitations in these experiments: Lord (2005) and Ausín and Sutton (2010) did not include control groups; improvement was measured differently, i.e., acoustically in Lord (2005), but with judgments in Elliot (1997) and Trofimovich and Gatbonton (2006). Also, the targeted modality used to gauge improvement was production for all studies except Ausín and Sutton (2010). However, studies such as Chela-Flores (2001), and Arteaga (2000) suggest that perception should precede production activities at the beginning level, with Arteaga (2000) advocating nontechnical linguistic explanations, and Chela-Flores (2001) proposing a more embedded and meaningful context for the instruction. The current study addresses these concerns by incorporating perceptual activities in a beginning L2 classroom, while using both technical and non-technical forms of instruction to determine its effects on learners.

The following research questions motivate the present study:

1. Is pronunciation instruction beneficial for the perceptual development of beginning Spanish learners?

2. Does the type of instruction (technical or non-technical) have an effect on learners' perceptual development and accuracy in Spanish?

3. Will beginning L2 learners' perception improve more on certain sound groups than others?

METHOD

Participants

The participants in this study were students in three sections of a second-semester Spanish class at Florida State University. The language history questionnaires (Appendix A) revealed that all participants were born in the United States and native speakers of English. No participants reported being bilingual or speaking a language other than English at home. The three sections were divided by type of instruction: one section (*n*= 19) received technical pronunciation instruction (TI); one section (*n*= 17) received non-technical pronunciation instruction (NI); one section (*n* = 17) received no pronunciation instruction and served as a control group (C). Both instructors were graduate students studying Spanish linguistics.

Treatment

The treatment started in the experimental sections after a pretest was administered to all sections. One topic was taught each week for ten minutes over ten weeks. A full schedule of topics can be found in Appendix B.

In the TI group, the target sounds were explained in terms of their manner and place of articulation with linguistic terms that were defined in class. Animated vocal tract diagrams[1] were also used. Participants were also encouraged to think about how these sounds compared with similar English sounds, and the instructor explained common pronunciation errors made by native English speakers. In the NI group, the instructor demonstrated the target sounds orally and then exemplified these sound in the contexts of real Spanish or English words. After the explanation of the target sounds, both experimental groups completed a perception activity that included listening to recordings of discrete Spanish words and answering questions based on what they heard. Types of activities usually had six to eight questions and included *picture identification*, choosing between a target and non-target like Spanish pronunciation, or *identifying* a speaker's regional dialect. The control group received neither instruction nor perception activities. All the groups (including the control group) received input from their normal class assignments, which included activities where learners must listen to and understand audible speech in order to answer questions.

Assessment

The participants' perception was tested with three perception tests: the pretest (administered before the treatment began), the immediate posttest (one week after completing the treatment), and the delayed posttest (four weeks after the treatment). Each test contained a total of 240 items divided into two blocks. Each test lasted about 20 minutes. Of the 240 total items, 80 were fillers. During the tests, the participants saw a written Spanish word for two seconds before hearing its pronunciation. There were no repetitions. Participants were then asked the target question, "*Was this word pronounced correctly*?" or the filler question, "*Have you heard this word before*?" Participants then used a response pad to press either 'yes' or 'no', with no time limit. It is important to mention that all items used were Spanish words selected from the textbook used in the participants' class (VanPatten, Leeser, & Keating, 2011). All items used were recorded by native Spanish speakers in both a target and non-target like condition (see Appendix C for a detailed description of the non-target like items), which were counterbalanced across blocks. The native Spanish speakers included both male and female graduate students from various regional dialects in order to provide a balance of genders, dialectal features, and also to mimic the variation present in the learners' normal homework assignments[2].

Data Analysis

Participants' accuracy means were analyzed with a 3 x 3 ANOVA with Group (TI, C, NI) as a between subjects variable, and Test (pretest, posttest 1, posttest 2) as the within subjects variable. Three participants who demonstrated performance at random in the tests (scores lower than 55%) were excluded from the analysis, as well as two participants who were not present for all assessments. There were 48 participants in total. (Note: Five participants in the TI group did not complete the posttests and are excluded from the analysis.)

---

[1](http://www.uiowa.edu/~acadtech/phonetics/spanish/frameset.html)

[2] All sections of this course had an on-line homework component, which included listening activities that were recorded by male and female speakers of various dialects.

RESULTS

Table 1 presents the descriptive statistics for each group and test. The ANOVA revealed no significant effect for group, $p = .259$. There was a significant main effect for test, $F (2, 90) = 65.3$, $p < .001$, with the highest perception accuracy percentage in the immediate posttest. There was also a test $\times$ group interaction that approached significance, $F (4, 90) = 2.4$, $p = .054$.

Table 1
Descriptive Statistics for Pretest and Posttests by Group

Group	N	Pretest		Posttest 1		Posttest 2	
		M	*SD*	*M*	*SD*	*M*	*SD*
1 (TI)	14	63.3	8.9	69.3	5.9	64.2	6.2
2 (C)	17	62.4	4.4	68.8	4.4	63.9	4.8
3 (NI)	17	64.1	4.6	75.1	5.7	64.3	6.1

The test $\times$ group interaction was analyzed using a test of simple main effects with Bonferroni adjustment for multiple comparisons. Pairwise comparisons for the interaction revealed that, on the immediate posttest, the non-technical group (NI) significantly outperformed the control (C), $p = .020$, and the technical groups (TI), $p = .050$. No significant differences were found between the groups in either the pretest or the delayed posttest. The ANOVA also revealed that all three groups made significant gains from pretest to immediate posttest, $p < .001$; however, the gains were not sustained in either group by the time the delayed posttest was administered, $p < .001$ (see Figure 1).
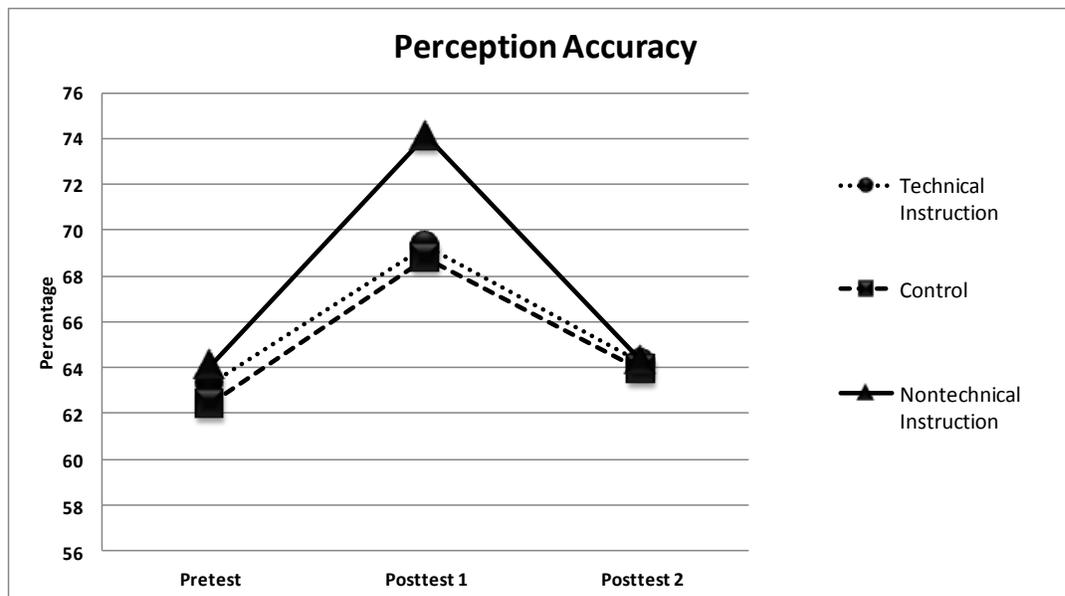

Figure 1. Accuracy in perception test by group.

A repeated-measures ANOVA with Sound (see Table 2 for the sound list), Condition (target or non-target like), and Test (pretest, posttest 1, posttest 2) was also performed to explore whether specific groups of sounds improved with the treatment. The analysis revealed a significant main effect for test, F (2, 90) = 76.49, $p < .001$, sound, F (11, 484) = 75.58, $p < .001$, and condition, F (1, 45) = 689.75, $p < .001$. There was no significant effect for group, $p = .367$. There were several significant interactions between sound, condition, test, and group, which are summarized in Appendix D. Relevant to this paper is the test × sound interaction, F (22, 968) = 16.29, $p < .001$, which was analyzed further with an analysis of simple main effects (Bonferroni adjustment).

Pairwise comparisons for the test and sound interaction showed significant improvements from pretest to immediate posttest for eight of the twelve sound groups tested. A summary is displayed in Table 2.

Table 2

Pairwise Comparisons from Pretest to Posttest by Sound

Sounds	Pretest *M*	Posttest 1 *M*	Mean Difference	*SE*	*P*
Vowels	67.8	58.1	-9.7	1.6	< .001
Diphthong/Hiatus	59.8	67.4	7.7	2.1	= .002
/p, t, k/	52.4	65.2	12.7	1.4	< .001
Vibrants[3]	59.1	59.6	0.6	1.4	= 1.000
Laterals	69.1	88.8	19.8	2.5	< .001
Nasals	72.3	78.1	5.7	1.9	= .016
'h', /h/	76.2	80.5	4.3	2.0	= .121
'z', /s, θ/	71.3	77.4	6.1	1.9	= .007
'v'	54.5	51.2	-3.3	2.5	= .059
[g, tʃ]	63.6	80.1	16.5	2.6	< .001
[β̞,ð,ɣ]	49.7	66.7	17.1	1.8	< .001
'y', 'll'	54.7	72.5	17.8	2.4	< .001

Pairwise comparison between immediate and delayed posttests revealed a different pattern. From immediate to delayed posttest, the groups showed significant gains on only two of the twelve sound groups tested: vowels and 'v'. There were significant losses on eight sound groups (see Table 3).

---

[3] The term *vibrants* refers to rhotics that are either realized as *taps* [ɾ] (simple vibrant) or *trills* [r] (multiple vibrant).

Table 3
Pairwise Comparisons from Immediate to Delayed Posttest by Sound

Sounds	Posttest 1 M	Posttest 2 M	Mean Difference	*SE*	*P*
Vowels	58.1	68.2	8.2	1.8	< .001
Diphthong/Hiatus	67.4	32.9	-5.6	2.0	= .025
/p, t, k/	65.2	58.6	-9.5	1.2	< .001
Vibrants	59.6	57.9	-1.8	1.6	= .834
Laterals	88.8	71.5	-17.3	2.4	< .001
Nasals	78.1	77.2	-0.8	1.7	= 1.000
'h', /h/	80.5	79.7	-1.4	2.2	= 1.000
'z', /s, θ/	77.4	71.8	-5.6	1.8	= .009
'v'	51.2	58.3	7.1	2.3	= .012
[g, tʃ]	80.1	62.8	-19.3	2.6	< .001
[β, ð, ɣ]	66.7	46.7	-20.1	1.7	< .001
'y', 'll'	72.5	54.2	-18.3	2.3	< .001

The analysis of learners' accuracy in perception also revealed significant gains from pretest to delayed posttest for /p, t, k/, *p* = .027, and a marginal effect was found for nasals, *p* = .058. All other sound groups showed no significant difference in accuracy between pretest and delayed posttest.

The two sound conditions within each of the tests were analyzed as well, and these results are shown in Table 4.

Table 4
Accuracy of Target and Non-target like Conditions by Tests

Test	Conditions			
	Target-like		Non-target like	
	M	*SD*	*M*	*SD*
Pretest	89.2	7.0	35.9	2.0
Immediate posttest	90.4	5.0	50.5	1.6
Delayed posttest	89.7	5.0	37.2	1.8

The results revealed that participants had significantly higher accuracy when identifying target-like conditions than non-target like conditions, $p < .001$, in all tests. Also, they demonstrated a significant improvement in identifying non-target like conditions from pretest to immediate posttest, $p < .001$. However, from immediate to delayed posttest, accuracy of non-target like conditions declined significantly, $p < .001$. This is displayed in Figure 2.
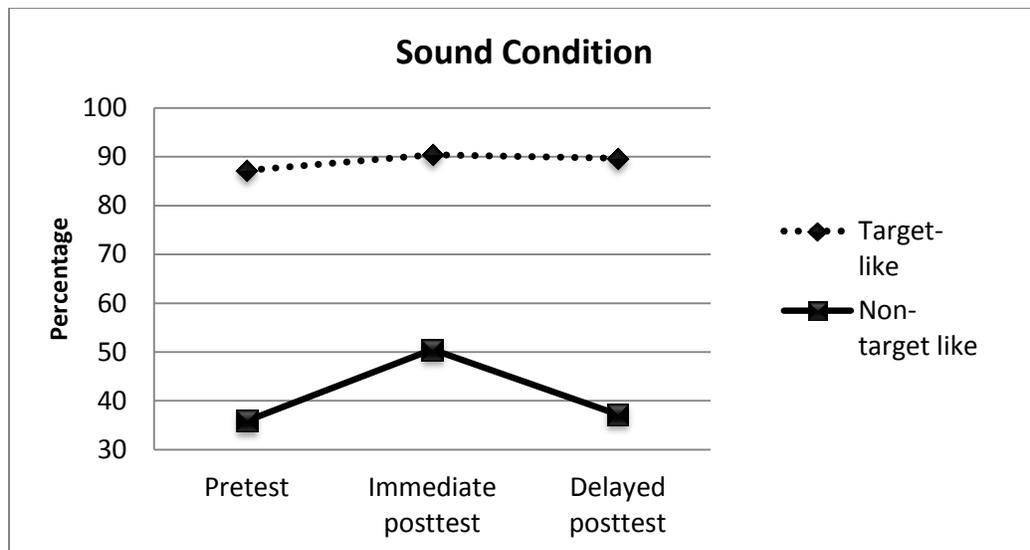

*Figure 2.*Accuracy in perception test by sound condition.

To summarize, the results suggest that all three groups improved from pretest to immediate posttest on their perception of non-target like sounds. On the immediate posttest, the NI group outperformed the TI and the Control groups. The Sound analysis revealed that there was significant improvement in the participants' perception of eight of the twelve target sounds examined from pretest to posttest (diphthong/hiatus, laterals, /p t k/, nasals, [g tʃ], [β̞ ð ɣ], 'z'/s θ/, and 'y''ll'). From pretest to delayed posttest, only two sound groups showed improvement (vowels, 'v'). The analysis also revealed that the improvement in perception was mainly in the identification of non-target like sounds compared to target-like sounds. The improvement in perception ability was not sustained by any of the groups by the time of the delayed posttest was administered.

DISCUSSION

In this study, the groups showed significant improvement on the perception task from pretest to immediate posttest, regardless of instruction. This seems to suggest that the type of input provided to the three groups during the language class was sufficient to draw the learners' attention to different types of sounds. It is possible that the input, along with these types of aural and meaning-based activities, helped the groups improve their perceptual abilities. However, the groups also showed significant losses from delayed to immediate post-tests, suggesting an effect for time for all three groups or a more universal U-shaped learning trend, in which the delayed post-test represents the valley of the skill in question.

An interesting finding is that the TI group performed just like the control group, and the NI group outperformed these two groups. These results suggests that non-technical instruction may be more beneficial in helping beginning learners improve their perception skills than technical

instruction. These results are consistent with the results in Trofimovich and Gatbonton (2006), who found that pronunciation activities that focused on *form* as well as in meaning were less successful for beginners than for higher proficiency learners. The combination of explicit phonetic forms as well as the normal meaning-based audio activities used in the TI group could have been cognitively more demanding for beginning learners who may still be in a 'meaning before non-meaning' stage of acquisition (Van Patten, 2007). On the other hand, reinforcing the normal meaning-based audio activities with more input (i.e. the contextual, non-technical examples) in the NI group could have strengthened the learners' perceptual skills by bringing their attention to the target sounds and helping them "notice" them, a construct which Schmidt's (1990) Noticing Hypothesis argues is necessary for acquisition. When the type of input was no longer available to the NI group, the improvement differential did not persist.

Another interesting finding was that although the groups showed significant gains on eight out of twelve target sounds from pretest to immediate posttest, their accuracy declined significantly from immediate to delayed posttest, resulting in sustained significant gains only for /p, t, k/, and a marginal effect for nasals, and significant immediate to delayed posttest gains for vowels and 'v'. Best and Tyler's (2007) Perceptual Assimilation Model for L2 learners (PAM-L2) posits that various factors affect the likelihood of a sound or sound group being acquired, such as L1-L2 distribution of the sound(s), frequency, or perceptual sensitivity. Several of these factors may have played a role for the overall improvement of the four sound groups that improved at the delayed post-tests. To explain, the nasal consonants /m/ and /n/, vowels, and (unaspirated) /p, t, k/ have similar distributions in both Spanish and English, making this a 'single-category' transfer process, which the PAM-L2 model considers the most optimal transfer situation. The majority of the sounds that improved at the delayed post-test are also frequent sounds in Spanish. As documented by Guirao and Jurado (1990), vowels (except /i/) and /n/ were amongst the 67.5% of all sounds measured, followed by the group that contained /m/, /i/, and the voiceless stops /p, t, k/. Perceptual sensitivity was also a factor for nasals, which an additional reaction times analysis revealed to be the fastest sound group perceived by participants in both conditions, and /p, t, k/, which was perceived faster in its target-like condition. As predicted by the PAM-L2 model, a combination of these factors seems to provide the most optimal environment for perceptual acquisition and transfer.

The accuracy of the target and non-target like conditions, as well as their reaction times on the tests, revealed that sounds in the target-like condition showed a ceiling effect (with an accuracy of approximately 90%) and were perceived significantly faster than sounds in the non-target like conditions. Each word in the non-target like condition consisted of L1 and L2 sounds; therefore, a possible explanation for these results is that the non-target condition was cognitively more costly to process because it activated the participants' L1 and L2. More research is needed in order to understand how processing load affects perception.

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

There are several limitations within this study. One is the grouping of target sounds; in cases such as the vibrants (taps and trills), where different perceptual processes may be taking place with simple and multiple vibrants, it might be more appropriate to separate sounds like the simple (tap) and multiple (trill) vibrants into different sound groups for analysis. The relatively small number of participants due to class attrition and outliers is also an issue; the availability of more participants could contribute to more robust, reliable findings.

Future studies could examine how perception or pronunciation instruction could relate to the phonological development of several levels of L2 learners and include delayed assessments. Due to the difficulty involved in incorporating a phonetics component into an established curriculum, perhaps a smaller group of sounds (those which showed the most improvement) could be considered for implementation. Learners in study abroad or naturalistic contexts could also be investigated, in which proficiency level will need to be determined by means other than course title. Studies that include both production and perception measures could attempt to explain the link and suggested gap that exists between perception and production, and inform pedagogical approaches about learner development and appropriate interventions.

CONCLUSION

The results from a pretest and two posttests measuring L2 perception of target and non-target like pronunciations of Spanish words suggest that explicit pronunciation instruction may not be beneficial for the perceptual development of second-semester learners. All groups involved in the study improved significantly from pretest to immediate posttest, suggesting the importance of meaningful input in the perceptual development of L2 learners, and there were no statistical difference between the scores of the control and explicit information group on the pretest or delayed posttest. However, the nontechnical instruction group did significantly outperform the control group and the technical on the immediate posttest, which suggests that non-technical linguistic descriptions and meaning-based forms of input may be more beneficial for beginning learners, although this effect was lost with time. Beginning L2 systems may not be ready to process the demands of tasks that focus both on meaning and explicit phonetic forms, but incorporating a less formal approach in these beginning courses may aid in learners' perceptual development.

ABOUT THE AUTHORS

Ann Aly Bailey is currently a doctoral student in the Department of Linguistics at the University of California, Los Angeles and received her M.A. in Spanish Linguistics from Florida State University. Her research interests include acquisition of Romance languages, phonetics and phonology, and historical sound change. Ann has previously researched the acoustics of Portuguese vowels by native and non-native speakers and hopes to research Caribbean varieties of Spanish in the future.

Anel Brandl teaches in the Department of Modern Languages and Linguistics at Florida State University. She specializes in second language acquisition and Spanish phonology, and she regularly teaches undergraduate and graduate courses in these areas. Her current research examines the processing strategies utilized by beginning second language learners of English and Spanish. She is also exploring the processing of inflectional morphology in beginning second language learners.

Contact Information

Ann Aly Bailey
University of California at Los Angeles
5310 Rolfe Hall, Box 951532
Los Angeles, CA 90095
ann.aly.bailey@gmail.com

Anel Brandl
Department of Modern Languages & Linguistics
The Florida State University
362 Diffenbaugh
Tallahassee, FL 32306
abrandl@fsu.edu

REFERENCES

Arteaga, D. L. (2000). Articulatory phonetics in the first-year Spanish classroom. *The Modern Language Journal*, *84*, 339-54.

Ausín, A., & Sutton, M. (2010). *An L2 pronunciation judgment task*. Paper presented at Selected Proceedings of the 12th Hispanic Linguistics Symposium, Somerville, MA: Cascadilla Proceedings Project.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O. S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Chela-Flores, B. (2001). Pronunciation and language learning: An integrative approach. *IRAL*, *39*, 85-101.

Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania, 80*, 95-108.

Guirao, M., & Jurado, M.A. (1990). Frequency of occurrence of phonemes in American Spanish, *Revue Quebecoise de Linguistique, 19*(2), 135-150.

Hurtado, L. M., & Estrada, C. (2010). Factors influencing the second language acquisition of Spanish vibrants. *Modern Language Journal, 94*, 74-86.

Lord, G. (2005). (How) can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania, 88*, 557-567.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129-158.

Swain, M. (1995). Three functions of output in second language learning. *Principle and practice in applied linguistics: Studies in honour of HG Widdowson*, (pp. 125-144). Cambridge: Cambridge University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams, (Eds). *Focus on form in classroom second language acquisition* (pp. 64-82). NY; Cambridge University Press.

Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *The Modern Language Journal*, *90*, 519-35.

VanPatten, B., & Williams J. (Eds.). (2007). *Theories in second language acquisition: An introduction*. Mahwah, NJ: Erlbaum.

VanPatten, B. (Ed). (2004). *Processing instruction. Theory, research and commentary*. Mahwah, NJ: Erlbaum.

VanPatten, B., Leeser, M., & G. Keating. (2011). *Sol y viento: Beginning Spanish* (3[rd]ed.). New York, NY: McGraw Hill.

Appendix A
Language History Questionnaire

Subject # _____

Date _____

This questionnaire is designed to give us a better understanding of your experience with other languages. We ask that you be as accurate and thorough as possible when answering the following questions.

General Background Questions

1. Gender
□ Female
□ Male

2. Age: _____ years

3. Do you have any known visual or hearing problems (corrected or uncorrected)?
□ No
□ Yes [Please explain] _____

4. Native Country
□ United States
□ Other _____
If other, at what age did you come to the US? _____

Home Language:
5. What is your native language?
□ English
□ Other: _____

6. Language spoken at home:
□ English
□ Spanish
□ Other _____

Education:
7. Please indicate where you have studied Spanish.
Please check all that apply and indicate length of study.

High School   College
□ 1 year   □ Less than a one semester
□ 2 years   □ 1-2 semesters
□ 3 years   □ 3-4 semesters
□ 4 years   □ 5-6 semesters
                                                 □ 7-8+ semesters

Rate your Spanish Skills:
8. Please rate your Spanish reading proficiency. (1=not literate and 10 = very literate)

1 2 3 4 5 6 7 8 9 10

9. Please rate your Spanish writing proficiency. (1=not literate and 10=very literate)

1   2   3   4   5   6   7   8   9   10

10. Please rate your Spanish speaking ability. (1=not fluent and 10=very fluent)

1   2   3   4   5   6   7   8   9   10

11. Please rate your Spanish speech comprehension ability. (1= unable to understand conversation and 10=perfectly able to understand)

1   2   3   4   5   6   7   8   9   10

12. Rate how comfortable you feel expressing yourself in Spanish:

1   2   3   4   5   6   7   8   9   10

13. Have you had any musical training (this includes informal experiences with singing and musical instruments)?
        NO_____
        YES How long? _____
        What kind of instrument or ensemble? _____

14. Is there anything else that we should know about your language abilities? Other languages you may speak, etc. Please explain:

_____

_____

_____

_____

_____

Appendix B
Schedule of Treatment

Practice #	Week	Date	Topic
1	2	1/11	Pretest
2	3	1/18	Simple and Multiple Vibrants
3	4	1/25	The letter 'h': Orthography vs. pronunciation
4	4	1/27	/p, t, k/: Orthography and Non-Aspiration
5	5	2/1	Spanish nasals
6	6	2/8	The letters 's, z, c': Dialectal variations
7	7	2/13	The letters 'g, j, ch'
8	7	2/15	Approximants and the letter 'v'
9	9	2/27	Laterals and the letter 'y': dialectal variations
10	9	2/29	Spanish vowels
11	11	3/14	Diphthongs vs. hiatus
12	12	3/21	Posttest
13	16	4/20	Delayed Posttest

Appendix C
Definitions of Non-targetlike Conditions

Sounds	Types of errors used and examples
Vibrants	The target vibrant was replaced by its counterpart (i.e. a multiple vibrant was replaced by a simple vibrant and the simple vibrant was replaced by a multiple vibrant). ○ Example: *barra* ('bar') ['ba.ra] pronounced as *['ba.ɾa]
Vowels	Vowels were replaced with an English-like vowel or diphthong ○ Examples: *café* ('coffee') [ka.'fe] as *[kæ.'fe]; *lejos* ('far') ['le.xos] as *['le.xoʷs]
Diphthong/Hiatus	Tokens with diphthong or hiatus were replaced with their counterpart, i.e. two vowels in hiatus were pronounced as a diphthong and a diphthong was pronounced as two vowels in hiatus. ○ Examples: *edificio* ('building') [e.ði.'fi.sjo] pronounced as * [e.ði.fi.'si.o]; *filosofía* ('philosophy') [fi.lo.so.'fi.a] as *[fi.lo.'so.fja]
Laterals	Word final /l/ was velarized and pronounced as [ɫ]. ○ Example: *fiel* ('faithful') [fjel] was pronounced as *[fjeɫ]
Voiceless occlusives	/p, t, k/ were pronounced with aspiration in non-target conditions. ○ Example: *pantalones* ('pants') [pan̪.ta.'lo.nes] was pronounced as*[pʰan̪.ta.'lo.nes]
Voiced occlusives	Intervocalic /b, d, ɡ / were not spirantized or were de-voiced in non-target conditions. ○ Examples: *lado* ('side') ['la.ðo] pronounced as *['la.do]; *conseguir* ('to get') [kon.se.'ɣiɾ] as *[kon.se.'kiɾ]
Nasals	The nasal consonants /n, m,ɲ/ were interchanged in the non-target conditions. ○ Examples: *mañana* ('morning','tomorrow') [ma.'ɲa.na] pronounced as: *[ma.'na.na]; *medias* ('socks', 'tights') ['me.ðjas] as *['ne.ðjas]
Affricate	/tʃ/ was pronounced as [ʃ] in non-target conditions. ○ Example: *salchicha* ('hot dog') [sal.'tʃi.tʃa] pronounced as *[sal.'ʃi.ʃa]
'z'	Orthographic 'z' was pronounced as [z] in non-target conditions ○ Example: *zanahoria* ('carrot') [sa.na.'o.ɾja] pronounced as *[za.na.'o.ɾja]

'v'	Orthographic /v/ was pronounced as [v] in non-target conditions.
	o Example: *viernes* ('Friday') ['bjeɾ.nes] pronounced as *['vjeɾ.nes]
'gui'	The orthography 'gui' was pronounced as [gwi] in non-target conditions.
	o Example: *guitarra*('guitar') [gi.'ta.ra] pronounced as *[gwi.'ta.ra]
'h'	Tokens with the orthographic letter 'h' were pronounced as /h/ in the non-target conditions; tokens with a phonetic /h/ represented by an orthographic 'j' or 'g' were pronounced as [dʒ].
	o Examples: *hora* ('hour') ['o.ɾa] pronounced as *['ho.ɾa]; *gemelos*('twins') [xe.'me.los] pronounced as *[dʒe.'me.los]
'll'	Tokens with orthographic 'll' were pronounced as /l/.
	o Example: *mantequilla* ('butter') [man̪.te.'ki.ʝa] pronounced as *[man̪.te.'ki.la]
'y'	Orthographic 'y' was pronounced as the glide [j] in non-target conditions.

Appendix D
ANOVA TABLE

Source	df	MS	F	p	η^2_p
Test × Group	4	.08	3.03	= .022	.12
Sound × Group	22	.07	2.36	= .001	.10
Test × Sound	22	.30	16.29	< .001	.27
Test × Condition	2	1.60	34.89	< .001	.44
Sound × Condition	11	2.63	81.85	< .001	.65
Sound × Condition × Group	22	.56	1.75	= .019	.07
Test × Sound × Condition	22	.11	6.50	< .001	.13
Error	968	.02			

ESL TEACHERS' BELIEFS AND PRACTICES IN PRONUNCIATION TEACHING: CONFIDENTLY RIGHT OR CONFIDENTLY WRONG?

Ron I. Thomson, Brock University

This study examines the extent to which English Language Teachers (ELTs) are able to critically evaluate beliefs and teaching practices in the domain of pronunciation instruction. To answer this question, an online survey was administered to 58 teachers, asking them to indicate to what extent they agreed or disagreed with statements taken from websites and *YouTube* videos belonging to pronunciation, accent reduction and accent modification programs. While respondents generally agreed with the least controversial statements in the survey, there was substantial disagreement about some of the more questionable or even false statements. Results indicate that many ELTs seem not to have the background knowledge, and lack the confidence necessary, to critically assess questionable pronunciation beliefs and practices – beliefs and practices that they may encounter in the materials that they choose to use. This suggests that more language teacher education programs need to offer courses in how to teach pronunciation. Furthermore, such courses should provide not only pedagogical techniques, but the theoretical knowledge necessary to understand how second language pronunciation develops. This will allow teachers to more effectively use commercial materials, while avoiding materials that are based on a faulty understanding of what works.

INTRODUCTION

Although pronunciation instruction was an important feature of many language teaching methods in the 20th century, the advent of Communicative Language Teaching (CLT) in the 1980s led many teachers to believe that a focus on form, including pronunciation, was unnecessary (Thomson, 2012a). Today, pronunciation instruction is experiencing resurgence. This is evident from the marked increase in the publication of materials on this topic, for both teachers and students (Foote, Holtby & Derwing, 2011), and from the popularity of pronunciation workshops at teachers' conferences (Thomson, 2012a). It is also apparent from a rapid growth in the number of individuals offering accent reduction or accent modification services (Thomson, 2013), which some have described as a lucrative and best-kept-secret career (Kuster, 2010; Nemko, 2008).

This renewed interest in pronunciation instruction has occurred during what Richards and Rogers (2001) describe as a post-methods era, characterized by an eclectic approach to language teaching, rather than strict adherence to any prescribed method. Thus, instead of there being a coherent and unified approach to teaching, as was characteristic of the pre-CLT era, there now exists a plethora of competing beliefs and practices, many of which are not evidence-based (see Thomson, 2012a). Coupled with an emerging but unregulated marketplace for stand-alone pronunciation programs (often referred to as 'accent reduction' or 'accent modification'), the dissemination of questionable beliefs and practices, especially on the Internet, is rapidly increasing (see Thomson, 2013).

The proliferation of unregulated stand-alone pronunciation programs has led some to debate who is most qualified to teach pronunciation (Derwing, 2008; Lippi-Green, 2012; Thomson, 2012b).

Derwing (2008) argues that English Language Teachers (ELTs) possess the best foundation, because they have a more holistic understanding of the process of second language (L2) acquisition, and are also more likely to have the cross-cultural communication skills necessary to effectively serve this population. This does not preclude the possibility that non-ELTs may also acquire appropriate training, but it seems to imply that such training should be broad-based, rather than comprising only a brief course focused only on pronunciation teaching.

Paradoxically, while ELTs may have the best educational foundation, it is also clear that many are unqualified to teach pronunciation. A decade ago, several surveys of ELTs in Canada, Britain and Australia investigated the training and beliefs of ELTs with respect to pronunciation teaching (Breitkreutz, Derwing, & Rossiter, 2001; Burgess & Spencer, 2000; Macdonald, 2002). While these studies clearly indicated a growing awareness of the importance of teaching pronunciation, some ELTs cited a lack of adequate training as a reason for avoiding pronunciation instruction in their language classes.

More encouraging, Breitkreutz et al's (2001) study also found that some ELTs in the Canadian context had taken university courses in how to teach pronunciation, and that classroom practices were beginning to reflect knowledge of recent pronunciation research, and advice found in reputable teacher resources. For example, many teachers were aware that accurate suprasegmental features of pronunciation (e.g., word stress, sentence stress, rhythm, intonation, etc) are considered more important to communication than are segmental features (i.e., vowels and consonants). Despite these positive findings, Breitkreutz et al. still called for better access to training in how to teach pronunciation.

Ten years after Breitkreutz et al's study, Foote et al. (2011) conducted a follow-up survey in order to determine if pronunciation teacher training opportunities had further advanced, and whether teaching practices by ELTs in the Canadian context had continued to improve. They found that little had changed. Particularly concerning was an indication that access to university-based credit courses dedicated to the topic of teaching pronunciation remained very limited. Instead, most ELTs continue to rely on presentations and workshops at teachers' conferences, where the expertise of those presenting is sometimes uncertain.

Given that professional development opportunities related to pronunciation teaching are limited, many ELTs are left to determine their own approach; reliance on commercially published and online sources is one obvious strategy. Thus, while the recent increase in pronunciation materials indicates progress, to provide appropriate instruction, teachers must be capable of discriminating materials and techniques that are evidence-based from those that are not. For example, Derwing (2008) reports that one 'accent reduction' program asks learners to recite "Peter picked a peck of pickled peppers" with a marshmallow placed between their lips. This technique seems to be based on the faulty assumption that lips need to be strengthened in order to produce a /p/. Even more unorthodox, J. Thompson's (2011) pronunciation text 'English is Stupid' indicates that the difference between /p/ and /b/ is that to produce the first sound one must breathe out, while to produce the second sound one must breathe in. While it is safe to assume that most ELTs would recognize how these particular suggestions conflict with what is known about English pronunciation, other debatable beliefs and techniques may not be as easy to recognize.

The purpose of this study is to investigate the extent to which ELTs are able to critically evaluate beliefs and practices about pronunciation instruction found on the Internet and in some self-published materials. If teachers have adequate preparation for teaching this skill, and can thereby

accurately assess the quality of materials they encounter, they should be less susceptible to adopting the inaccurate beliefs and practices that are increasingly prevalent in the marketplace.

METHODS

Survey instrument

The survey largely followed the format used by Foote et al. (2011), and contained some of the same demographic questions. However, most items were new, and required respondents to evaluate belief statements or techniques taken verbatim from pronunciation teaching materials found on websites and in *YouTube* videos promoting pronunciation, accent reduction or accent modification services. In a few cases, statements were taken from self-published texts. The 24 items reported in this study are a subset of a much larger survey, which included 131 items related to pronunciation beliefs and practices[1]. While a few open-ended questions were used, the majority of questions were either multiple-choice or asked participants to indicate the degree to which they agreed with particular statements. The entire larger survey took approximately 45 minutes to complete.

Participants

To recruit respondents, I posted invitations to participate in relevant pronunciation interest groups on the social media site LinkedIn. I also asked colleagues at Canadian English language programs to forward my invitation to teachers in their institutions.

RESULTS

A total of 58 ELTs responded to the survey, including 43 from Canada and 15 from the United States (12 male; 46 female). Most (86%) were native speakers, with the rest having a variety of L1 backgrounds.

The majority of participants had a Master's (62%) or undergraduate degree (30%) in TESL or Linguistics, while the rest (8%) had a degree in an unrelated field. Nearly a quarter (23%) reported having taken an entire course in pronunciation teaching, while 66% reported attending workshops and presentations on the topic at teachers' conferences. Although all indicated that they had taught pronunciation, only 75% reported feeling qualified to do so; 67% indicated a desire for more training in this area.

Respondents were asked to respond to a series of statements in four broad categories: 1) the nature of a foreign accent, 2) general instructional strategies, 3) descriptions of English sounds, and 4) specific teaching techniques. Although the respondents indicated their agreement using the labels 'strongly agree', 'agree', 'disagree', 'strongly disagree', and 'unsure', for ease of presentation, results are collapsed into three categories: 'agree', 'disagree' and 'unsure'.

Teachers generally agreed with the least controversial statements regarding the nature of a foreign accent (see Table 1). For example, there was strong consensus (93%) that transfer from a learner's first language is a significant cause of foreign accent. Similarly, most teachers (96%) agreed that learning to perceive sounds is important for developing accurate pronunciation.

The responses to the more questionable statements were more varied. A total of 75% agreed that pronunciation errors are due to weak speech muscles, while 41% agreed that a foreign accent

---

[1] The larger survey investigated the educational background, beliefs and practices of ELTs and Speech Language Pathologists who have an interest in pronunciation instruction.

stems from improper airflow. Most of the remaining teachers indicated that they were unsure about the accuracy of these statements. While very few teachers agreed that a foreign accent is like a communication disorder, nearly half were unsure if this was the case.

Table 1
Percentage agreement with statements regarding the nature of a foreign accent.

Statement	Agree	Disagree	Unsure
1. Accents are caused by carrying over the sound systems from students' native languages to their second language.	93	--	7
2. A key to good pronunciation is learning to hear the sound accurately.	96	2	2
3. Errors in pronunciation result from not having speech muscles that are properly toned for English sounds.	75	6	19
4. Improper airflow is a common cause of a foreign accent.	41	15	44
5. A foreign accent is not unlike other communication disorders	3	50	47

Table 2
Percentage agreement with statements regarding general instructional strategies.

Statement	Agree	Disagree	Unsure
1. Recording your voice allows you to monitor how you are producing your sounds.	88	2	10
2. Using a mirror to watch your mouth will help train your brain in a new habit.	66	6	28
3. Observing the mouth movements of NSs and trying to imitate them improves pronunciation.	61	10	29
4. Comparative sounds, alliteration, and tongue twisters can be used as drills to help improve your pronunciation.	77	5	17
5. Stretching sounds out will help learners to feel how their lips and tongue are supposed to feel while producing the sound.	62	11	27
6. Exercises will give you more control over your articulators so that you can make the sounds required in North American speech.	50	19	31
7. Some English sounds are difficult or impossible to create until you teach your speech muscles to relax and move in a particular way.	71	9	20

The first five statements in Table 2 relate to general strategies for raising learners' awareness about pronunciation. While most ELTs (71%) agreed with these relatively common practices, a large number (22%) were unsure. The final two statements in the table reflect a belief that exercising speech articulators is necessary in order to improve pronunciation. Again, the majority of teachers (61%) agreed, while many others (26%) were unsure.

Statements in Table 3 reflect ways in which some teachers describe specific English sounds to students. On average, only 14% agreed with these statements, while 35% were unsure. Although some of these statements are clearly difficult to interpret, some are not. For example, the first statement is unquestionably false. To know this, one only needs to attempt to follow this advice. Still, 10% agreed with the statement, while 14% were unsure.

Table 3
Percentage agreement with statements regarding descriptions of English sounds.

Statement	Agree	Disagree	Unsure
1. The difference between /p/ and /b/ is that for the first sound you breathe out, while for the second sound, you breathe in.	10	76	14
2. The difference between /ej/ and /ɛ/ is that one is tense and the other is relaxed or neutral.	26	17	57
3. To say the /ae/ sound keep your chin falling down and then bring it up to say an /ɪ/, the combination of these two sounds is the American /ae/ sound.	9	32	59
4. Just open your mouth to say the /a/ sound. It's like you are afraid "ahhhh".	24	40	36
5. The /r/ sound comes from your stomach. Your stomach moves in and you can feel it in your stomach.	14	62	24
6. When you say a word all you need to focus on is hitting and smashing the beginning. So, in this kind of speaking you have no syllables, in this language of accent reduction there are no syllables, in American English there are no syllables.	2	77	21

Regarding specific techniques that are recommended to improve pronunciation (see Table 4), there was also substantial disagreement. The first three statements recommend activities aimed at raising learners' awareness of articulation. Many disagreed with these techniques (31%), and even more (45%) were unsure. While most teachers also disagreed with the statement suggesting that students should practice speaking with marbles in their mouth, a large percentage (34%) were unsure if this was a good technique.

In response to earlier statements, the majority of respondents seemed to share a belief that weak speech muscles are a cause of mispronunciation. Yet, when it comes to what activities are useful for strengthening those muscles, there was greater disagreement. Most either disagreed (33%) or were unsure (44%) about the techniques suggested in the fifth and sixth statements in Table 4.

Table 4
Percentage agreement with statements regarding specific techniques.

Statement	Agree	Disagree	Unsure
1. Practicing speaking with a pencil in your mouth will help direct your attention to your articulators.	23	34	43
2. Stick your fingers in your mouth when you say the /r/ sound to help you feel the difference, as well as to keep you from lifting your tongue up.	11	44	45
3. Instruments placed in your mouth that position the tongue correctly can be used to correctly pronounce words with an American Accent.	12	41	47
4. Practice speaking with marbles in your mouth while reading. Take out the marbles and you will speak clearly.	4	62	34
5. Reading aloud in English for 15 to 20 minutes each day can improve your accent by strengthening the muscles of your mouth.	36	31	33
6. Strengthen your tongue by placing it on the roof of your mouth apply suction, and release making a popping sound.	16	29	55

When responses to statements are grouped in terms of those that seem least controversial (Table 1, statements 1-3; and Table 2 statements 1-4), and those that seem somewhat or very controversial (Table 1, statements 4 & 5; Table 2, statements 5-7; and all of Tables 3 & 4), an interesting pattern emerges. For the least controversial statements, teachers agreed in all instances (mean agreement 77%). Conversely, for the more questionable or unclear statements, teachers disagreed in only 29% of cases (mean disagreement 38%), although they were also more likely to be unsure (31%).

When responses made by ELTs who had taken an entire university credit course in teaching pronunciation were isolated from the rest of the sample, there was no evidence that they were better equipped to recognize the more questionable statements. For example, 57% of these more qualified teachers agreed that pronunciation errors result from having speech muscles that are not properly toned, while 36% agreed that improper airflow caused a foreign accent. A total of 14% agreed that learners can strengthen their tongue by placing it on the roof of their mouth applying suction and releasing it to make a popping sound; 14 % also agreed that an English /b/ is produced while breathing in.

DISCUSSION

The results of this survey reveal that ELT teachers who provide pronunciation instruction do not share a cohesive understanding of 1) the source of foreign accent, 2) general instructional strategies, 3) descriptions of English sounds, or 4) appropriate instructional techniques. Admittedly, some of the statements teachers were asked to evaluate are difficult to interpret, and may explain some of the disagreement found. Responses to other statements, however, suggest that many teachers' lack the ability to critically evaluate faulty beliefs and practices.

Agreement regarding the nature of a foreign accent

Most teachers were aware of the role that an L1 plays in pronunciation difficulties. They were also aware of the importance of developing speech perception. What is surprising is that agreement with the first two statements regarding the nature of accent was not unanimous - several teachers were unsure L1 transfer and speech perception played a role in pronunciation. This suggests that some practicing teachers lack even the most basic knowledge about pronunciation.

The fact that the majority of ELTs agreed that inaccurate pronunciation stemmed from having weak speech muscles, and that almost half also agreed that foreign accent is caused by improper airflow, is alarming. It seems unlikely that such knowledge would have been gained from most university courses dedicated to pronunciation teaching, since such topics are not discussed in popular pronunciation teacher training texts (e.g., Celce-Murcia, Brinton, Goodwin & Griner, 2010). However, this type of description is found in Speech Language Pathology (SLP) settings, where motor speech disorders are commonplace, and where articulation and breathing exercises are used in treatment (see Duffy, 2005). Only half of the ELTs were certain that a foreign accent was not a speech disorder, providing evidence that they may be influenced by beliefs and practices from the SLP profession. In fact, SLPs are the most dominant source of accent reduction and accent modification materials on the internet (see Thomson, 2013).

Contrary to the assertion that muscle weakness is a cause of foreign accent, many experts agree that the primary influences on foreign accent are age, L1-L2 perceptual similarity, and experience (Piske, MacKay, & Flege, 2001). Furthermore, Piske et al's exhaustive literature review does not even mention muscle weakness or airflow as contributing factors. Flege's (1995) Speech Learning Model, the most dominant model in the field of L2 speech learning, argues that improvement in perception will lead to improvement in production. R. I. Thomson (2011) demonstrated that at least for vowels, perceptual training can result in improvement in production, without articulation practice. The centrality of perception does not negate the value of production practice, since developing articulatory control and automaticity for L2 speech gestures is also an essential component of pronunciation. However, it is quite clear that emphasizing production practice over perceptual training is unjustified.

Agreement regarding general instructional strategies

Although most teachers agreed with statements regarding strategies for raising learners' awareness, many were still unsure. Since most of these statements were relatively uncontroversial, it is concerning that teachers lack confidence in agreeing that these practices are likely to help students. ELTs with formal education in the field should have some recognition that raising awareness of any linguistic forms (of which pronunciation is one) is an important precursor to acquisition (Schmidt, 2001).

Agreement regarding the description of English sounds

A small minority of teachers agreed with each of the statements concerning how English sounds are produced. This is an encouraging sign, since even relying on intuition should tell most teachers that the statements are unclear and unlikely to benefit learners. Some are even patently false. On the other hand, the fact that some teachers did agree with these statements, including those that are most absurd (e.g., breathing in to produce /b/), suggests that analytic skills may not be a strong suit of some teachers.

Agreement regarding specific instructional techniques

As with the description of English sounds, most teachers did not agree with the series of statements concerning specific instructional techniques. Again, appeal to intuition ought to rule most of them out as recommendable forms of instruction. For example, sticking their fingers or marbles in their mouths will make it very difficult to articulate the sounds that students are instructed to notice. Such practices are also unhygienic and even dangerous.

IMPLICATIONS

While previous studies have sought to uncover ELTs knowledge of pronunciation instruction and their teaching practices (eg., Breitkrutz et al., 2001; Foote et al., 2011), this study reveals the extent to which ELTs can critically evaluate beliefs and practices used by some in the field. On the one hand, it appears from earlier surveys that many ELTs are capable of remembering specific facts that they have been taught about pronunciation, and are able to apply that knowledge. On the other hand, this survey indicates that many ELTs, including those with specific training in pronunciation teaching, are unable to critically evaluate beliefs and practices they have not previously encountered. This makes them susceptible to following dubious advice found on the Internet and in other materials.

The most obvious implication is that TESOL/TESL programs need to do a better job of providing teachers not only with pedagogical knowledge, but also with critical thinking skills. While there may not be a 'correct' response to some of the statements found in this survey, ELTs should at least be able to critically evaluate such statements. They should not be confidently wrong in agreeing with statements that are clearly inaccurate. In contrast, the degree of uncertainty expressed by many of the survey's respondents can be seen in a positive light. It suggests that some teachers are aware that they lack the knowledge necessary to assess the validity of some beliefs and practices.

Universities that offer specific courses in teaching pronunciation should advise students to avoid materials taken from unpublished sources, unless they have strong evidence for the validity of those materials. English language programs can insure teachers are prepared by providing a library of legitimate, peer-reviewed pronunciation teaching resources, so instructors are not left to search on the Internet.

Finally, while I concur with Derwing (2008) that ELTs have the most appropriate foundation to provide pronunciation instruction, much progress is needed if they are to legitimately lay claim to this area of specialization. As disconcerting as some of their responses are, the ELTs in this survey likely reflect the very best of what the profession has to offer; since they self-selected to participate in the survey, they are more likely to have a particular interest in pronunciation than the ELT population at large.

ACKNOWLEDGEMENTS

ABOUT THE AUTHOR

Ron Thomson is an Associate Professor of Applied Linguistics at Brock University, in Niagara, Canada. He teaches in the undergraduate program in Applied Linguistics, and in the MA TESL

program. His research interests include oral fluency and pronunciation development, oral fluency instruction, high variability phonetic training, and computer assisted pronunciation teaching.

Ron I. Thomson
Department of Applied Linguistics, Brock University
500 Glenridge Ave
St. Catharines, ON, L2S 3A1
Canada

email: rthomson@brocku.ca
tel: 905-688-5550 ext 5842
fax: 905-688-2360

REFERENCES

Breitkreutz, J., Derwing, T.M. & Rossiter, M.J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal, 19*, 51-61.

Burgess, J. & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System, 28, 191-215*

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). New York: Cambridge University Press.

Derwing, T. M. (2008). Curriculum issues in teaching pronunciation. In J. G. Hansen Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 347-369). Philadelphia, PA: John Benjamins.

Duffy, J. R. (2005). *Motor speech disorders* (2nd ed). New York: Elsevier.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (Ed), *Speech perception and linguistic experience: theoretical and methodological issues* (pp. 229-273). Timonium, MD: York Press.

Foote, J. A., Holtby, A. K, & Derwing, T. M. (2011). Survey of pronunciation teaching in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29*(1), 1-22.

Kuster, J. M. (2010). Accent modification cited as "Best-Kept Secret". *The ASHA Leader.* Retrieved November 15, 2012 from www.asha.org/Publications/leader/2010/100427

Lippi-Green, R. (2012). *English with an accent: Language, ideology, and discrimination in the United States.* 2nd Edition. London: Routledge.

MacDonald, S. (2002). Pronunciation—views and practices of reluctant teachers. *Prospect, 17* (3), 3-18.

Nemko, M. (2008). Best-Kept-Secret Career: Accent-Reduction Specialist. *U.S. News & World Report*, Retrieved November 15, 2012 from http://money.usnews.com/money/careers/articles/2008/12/11/best-kept-secret-career-accent-reduction-specialist

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191-215.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Thompson, J. (2011). *English is stupid*. Caledon, ON: Thompson Language Center.

Thomson, R. I. (2011). Computer Assisted Pronunciation Training: Targeting second language vowel perception improves pronunciation. *CALICO Journal*, *28,* 744 -765.

Thomson, R. I. (2012a). Demystifying pronunciation research to inform practice. In McGarrell, H.M. & Courchêne, R. (Eds.) *Special Research Symposium Issue of CONTACT, 38*, (2), 63-75.

Thomson, R. I. (2012). Accent reduction. In C. A. Chapelle. (ed). *The Encyclopedia of applied linguistics.*(pp. 8-11). Oxford: Wiley Blackwell.
DOI: 10.1002/9781405198431.wbeal0004

Thomson, R. I. (2013, to appear) Accent reduction and pronunciation instruction are the same thing. In L. Grant (Ed.). *Pronunciation myths: Applying second language research to classroom teaching.* Ann Arbor: University of Michigan Press.

PRELIMINARIES TO HAPTIC-INTEGRATED PRONUNCIATION INSTRUCTION

William Acton, Trinity Western University
Amanda Baker, University of Wollongong
Michael Burri, University of Wollongong
Brian Teaman, Osaka Women's University

This paper reports on aspects of a haptic (movement plus touch) integrated system for classroom pronunciation instruction. It is based, in part, on established pedagogical practice in the use of somatic/kinesthetic techniques such as gesture in language instruction (Acton, 1984, 2012; Celce-Murcia, Brinton, Goodwin & Briner, 2010; McCafferty, 2004), and management of vocal resonance in singing and voice training (Lessac, 1997). The pedagogical method is designed for use by relatively untrained instructors and is generally best delivered through video with classroom follow up. Relatively recent research and development in haptics, especially in the areas of gaming, prosthetics and robotics, provides a rich source of potential principles and procedures from which to draw in exploring and rethinking the "clinical side" of pronunciation work. The use of haptic integration procedures in various teaching systems, in the form of designated movement patterns accompanied by various "textures of touch" has been shown to more systematically coordinate sensory modalities involved and greatly enhance both effectiveness and pace of instruction. In field testing the basic English pronunciation system to be described, this application of haptic procedures shows promise of also enhancing efficiency in anchoring sounds, words and phrases and in facilitating both recall and integration of targeted material in spontaneous speech.

PRELIMINARIES

The essential, haptic-integrated English pronunciation (EHIEP) framework (Acton, 2012; in preparation) outlined below involves a set of about a dozen protocols, or ordered sets of procedures, that train learners in haptic-based techniques for classroom instruction. "Haptic" in this context refers to systematic hand movement across the visual field accompanying speech that typically terminates in *a touch of some kind*, like one hand touching the other. That touch occurs simultaneously with the articulation of a stressed syllable of a word, focal stress of a phrase or a prominent word in discourse (e.g., Acton 2010; 2012, Acton, Baker & Burri, 2009; Burri & Rauser, 2010; 2011.) EHIEP is based on the broader notion of "haptic-integrated clinical pronunciation research" (HICPR, 2012) that focuses more explicitly on both the physicality of pronunciation and a clinical perspective on integration of enhanced pronunciation into spontaneous speech (c.f., ICPLA, 2013), both areas not generally addressed adequately in contemporary pronunciation teaching.

"Haptic-integrated" should be understood in two senses. First, these pedagogical movement patterns (PMPs) are used in the classroom in integrated instruction for various functions, including presenting, correcting, practicing or anchoring new or "repaired" sound structures. Second, the haptic-based PMPs serve to better enable recall and subsequent integration into spontaneous speech of targeted structures or processes.

The four sections of the paper correspond roughly to the contributions of each of the authors to the development of the system and their part of the oral presentation at the conference. The first, by Baker, reviews key research studies on integration of pronunciation, including both integrated instruction and integration of focus-on-form into spontaneous speech. The second, by Acton, addresses the application of embodiment theory, kinesthetic learning and haptic research to the problem of integration of pronunciation. The third, by Teaman, reviews the range of structures and processes which can be worked with effectively in the EHIEP system. The fourth, by Burri, then situates some of the methodology with typical anecdotal reports from field tests. (In the oral presentation, during Burri's comments, many of the PMPs and related techniques were demonstrated. Representative examples are linked on the HICPR blog (HICPR, 2012.)

INTEGRATION OF PRONUNCIATION TEACHING AND PRONUNCIATION

The field of second language teaching has seen a dramatic shift in how pronunciation has been situated in the classroom. It has moved from being an integral component during the Audiolingual period to a relatively minor sub-skill in the Communicative era and beyond (e.g., Brown, 2008; Derwing & Munro, 2005; Levis & Grant, 2003; Murphy, 1991; Naiman, 1992). The last decade has seen a shift in perspective with the emergence of two issues in particular: how to effectively integrate pronunciation into general instruction and how to facilitate the transfer of new and "changed" targets from classroom to conversation.

Effective integration of pronunciation instruction

Since at least the1980s, methodologists have advocated an integrated approach to pronunciation teaching (e.g., Celce-Murcia et al., 2010; Chela-Flores, 2001; Gilbert, 1987; Murphy, 1991). Integration can take different forms. Many consider integration to refer to any form of pronunciation instruction added to the curriculum, even on an ad hoc basis. Scales, Wennerstrom, Richard and Wu (2006), however, argued persuasively that pronunciation must instead be integrated into oral skills curricula.

There are any number of reasons why pronunciation instruction may not be addressed systematically. One difficulty is that there may not be sufficient time to teach all of what appears to be the relevant features (Munro & Derwing, 2006). To address this concern, Morley (1991), Hewings (1995), Celce-Murcia et al. (2010) and Acton (2012) have all highlighted the need for a "targeted" approach that focuses on intelligibility. At this point, from a research perspective, relatively little is known about how teachers do integrate pronunciation in their classrooms. One study, Baker (2012), examined the case of ESL teachers who, to varying extents, attempt to integrate pronunciation in a fairly fixed intermediate-level curriculum—with mixed results--and the questions that they have about how best to do that.

Transferring classroom change to the real world

The other perhaps more critical question is how to anchor changed pronunciation such that it leads to use in authentic contexts outside of the classroom. Here, too, research has yet to systematically examine the impact of instruction on use outside the classroom, in natural

settings. Nevertheless, there are a number of studies on changed pronunciation within the classroom or classroom-like contexts. Quasi-experimental research has demonstrated that explicit pronunciation instruction can have at least a short-term positive effect on learner pronunciation (Couper, 2003; Derwing, Munro & Wiebe, 1997; Macdonald, Yule & Powers, 1994; Saito & Lyster, 2012).

Those studies provide more than credible "evidence" that pronunciation instruction can impact acquisition of intelligible English. However, it is still some conceptual and theoretical "distance" from mimicking a target form or reproducing it adequately in controlled or guided activities in the classroom to using it at Starbucks. As of yet, few studies have adequately assessed the transfer of pronunciation instruction to contexts outside of the lab or classroom. What we do have are a few anecdotal reports from learners of decontextualized pronunciation practice and consequent difficulties with applying oral communication skills (e.g., Liu, 2012). So, how can we best teach pronunciation in the classroom so that it will appear as enhanced intelligibility "out there?" One possible approach is to first "embody" it.

EMBODIMENT AND HAPTIC INTEGRATION OF PRONUNCIATION

The EHIEP system takes a decidedly "embodied cognition" (McNerney, 2011) approach to integrated pronunciation instruction. That is, the role of the body in shaping mind and conceptual development must be considerably more foregrounded in our understanding of how we learn and most effectively teach pronunciation. Perhaps that not always to the extent that Lessac (1997) maintained, that we should "Train the body first!" in developing the voice and acting body, but at least there should be a balanced "felt sense" (Gendlin 1996) of cognitive and somatic awareness and engagement.

When we refer to our use of movement and touch we will use the term "haptic," as distinct from "haptics." "Haptics" now generally refers to sensing and manipulation through touch, related to its application in technological systems for virtual reality, gaming, training and rehabilitation—and, to some degree, smart phones. Our use of "haptic" is closer to that common in the arts, for example, haptic cinema, theatre, video or live performance—the idea being to convince the viewer's brain that "real" touch is involved, coming in through the eyes—without use of a piece of haptic hardware such as a glove, arm prosthetic or body suit—or water mist in the air or gyrating chair. A good introduction to the concept of "haptic" as it is embodied in media is Marks (2002).

In an important recent article, Holme (2012) situates embodiment theory in the field of English language teaching from a cognitive linguistics perspective, pointing to four general principles that are converging to rapidly redefine the field: conceptualization, the lexico-grammatical continuum, usage, and *embodied learning*. Haptic-integrated clinical pronunciation (HICP) (and EHIEP) generally fits within that last category, connecting perceived vocal resonance to targeted sounds by employing upper body movement and touch.

In pronunciation teaching the value of gesture and kinesthetic techniques has always been assumed (e.g., Celce-Murcia et al., 2010), but there has been little systematic instruction or study to date of the use of directed movement in the classroom or its impact back "on the street." There has been, on the other hand, substantial research (e.g., McCafferty, 2004) on the place of gesture in second language acquisition, in general, and related cross-cultural communication (e.g., Kendon, 2004).

In the earliest accounts of pronunciation teaching methods we find informal suggestions for gestural reinforcement of stress patterns, rhythm, intonation and other features (Acton, 1994; 1997). There are, as well, literally hundreds of teaching tips on the Web that allude to some use of the body in phonics and pronunciation instruction, for example, "kinesthetic phonics" and the well-known Total Physical Response teaching approach (Asher, 1972). For instance, Acton's (1984) article on dealing with fossilized pronunciation included reference to three kinesthetic techniques: using upper torso nods for practice phrasal stress, baton-like movement to control spontaneous speaking and using rubber band stretching to anchor vowel length, a commonly used technique today (Gilbert, 2012.)

In 2004 Acton began working on a research project with a team of counseling psychologists that involved the use of eye movement in the visual field, a therapy termed Observed Experiential Integration (OEI, 2012; Acton, in press). The focus of those therapeutic techniques is, in essence, rapid integration of changed attitudes and behaviors—achieved by working "through the body" to anchor change. Shortly thereafter, we began experimenting with similar strategies in pronunciation teaching. (For more specifics on the development of the EHIEP system and related research, see the HICPR blog, HICPR, 2012.)

Initially, the basic innovation of the EHIEP approach was to take commonly used classroom gestures and "anchor" them with touch, in effect making the use of those pedagogical movement patterns (PMPs) more consistent, systematic—and effective. The key elements of the techniques involved are:

- PMPs involve hands moving across the visual field, not unlike sign language.

- Locations in the visual field are associated with specific sounds or processes (related to OEI and other psychological models); high or low pitch, for example, may be associated with analogous relative positioning, up or down.

- Hands typically touch each other or one hand contacts a point on the body coinciding with lexical stress or discourse prominence. (That practice is also quite similar to some American sign language anchoring of highly emotional signs w/touch.)

- The PMPs appear to promote integration of targeted pronunciation. Although we lack "hard" evidence of that, other than a series of field tests in classrooms and clinical, one-on-one observations , our experience and the extensive research based on the use of haptic and haptics systems in general seems unequivocal in pointing to the power of haptic engagement to facilitate high-impact integrated learning.

EHIEP is designed for use by relatively untrained instructors, taught, if necessary, almost entirely through short videos which train learners and instructors in techniques for haptically anchoring and integrating essential features of basic English pronunciation.

ESSENTIAL ENGLISH PRONUNCIATION: LINGUISTIC PRELIMINARIES

This section describes the phonological categories that EHIEP emphasizes. The inventory of phonological targets intersects reasonably well with those used in general pronunciation instruction (e.g., Celce-Murcia et al., 2010; Gilbert, 2012). Even though EHIEP is quite different methodologically from standard approaches, the structures it deals with and its pedagogical

priorities should be familiar to most instructors. Although the approach is somewhat more holistic in its assumptions about the learning of pronunciation, it is generally consistent with current SLA theory as it relates to acquisition of the sound system. It is important to note, however, which features are explicitly included and which are not. So, before reviewing the structures that are covered, some comments will be made about what is not covered and why.

EHIEP does not attempt a comprehensive treatment of all phonological structures of English but instead embodies core parts of the L2 learner-oriented phonology—although, in principle, any phonological element or process is amenable to haptic embodiment for classroom use. Likewise, these principles could be applied to the sound system of any language.

EHIEP does not deal explicitly with many formal phonological rules, such as those relating to schwa or unaccented vowel qualities, for three reasons. First, for general intelligibility, some are simply not necessary or belong further along in the process. The overall priorities are close to those proposed by Jenkins (2000) in defining a general pedagogical model for "English as an International Language (EIL), as will be elaborated in what follows. Second, as noted earlier, EHIEP is based on the assumption that somatically grounded "exemplars" (c.f., Pierrehumbert, 2001) are key to learning the sound system. So, in the EHIEP approach, haptically reinforced,, context-situated lexical items are the primary targets of instruction. That does not exclude or restrict the use of more metacognitive explanations and rules; it only requires that sounds, words and phrases first be experienced as sensually, resonantly—and memorably as possible. Third, because the system is designed for instructors with little or no background in pronunciation teaching, much of the work must take place somewhat more experientially and inductively.

EHIEP does work extensively with the "feel" or felt sense of word and sentence stress in the body, but it does not explicitly teach word stress rules (such as methods to determine where word stress falls in individual words or compounds), spelling rules, phonics or formal phonetics. (Use of a standard, learner's dictionary with audio files is highly recommended.) In principle, although the sound system of any language can be haptically anchored, our work until now has focused on North American English (NAE) and the basic vowel systems of a few other languages.

In the following, the linguistic structures that are covered are described briefly in terms of how they relate to this approach.

Vowels

Standard vowels are covered: tense, lax and diphthongs. The inventory is consistent with most contemporary sources such as Celce-Murcia et al. (2010), which uses the terms "simple vowels," "vowels with an adjacent glide," and "diphthongs" to cover the same categories. Although the vowel coverage is for general NAE, other English vowel systems and dialects can be easily adapted. Each vowel is represented by a unique PMP that embodies it and is used in the learning process whenever that vowel in a word is targeted, typically in a stressed syllable.

Stress

For word and phrasal stress, one stressed syllable is targeted in an individual word or phrase. The focus is on the stressed syllable, itself, without paying undue attention to the unstressed syllables. (There is a PMP for secondary stress but it is not typically used when the focus is on simple intelligibility.) Word and phrasal stress are anchored and practiced in a fashion similar to that of vowels in that a PMP anchors the vowel in targeted stressed syllables. In general, stress assignment and anchoring (with PMPs) is only done in context, in conversational settings and

style, with texts to be spoken aloud, not read silently. No attention should be paid to complicated rules of stress, but simply to noticing and performing the stressed targets. That does not preclude, however, attention to such processes in related receptive skills and grammatical training.

Intonation

EHIEP uses three basic intonation contours (level, rise, fall) and two compound contours (rise fall, fall rise), which are made up of the three basic contours. An additional final fall (for discourse final utterances) or a final rise (often used to signal uncertainty) can also be attached to the end of a phrase. There are, of course, many other nuances to the intonation system, but that set seems to get learners connected to the important, basic forms, providing a solid basis for future work. (An advanced, "Expressiveness" module contains additional intonation-based features that are more discourse dependent.)

Rhythm grouping

Groups of one to seven syllables with one prominent syllable, are targeted and practiced. The rationale for the maximum number of syllables allowed is based, at least in part on research and experience in the limits of learner production groups. (Up through intermediate level, seven syllables is about all learners can manage without taking a break!) That involves identifying and haptically anchoring (with PMPs) the strongest, stressed syllable in each rhythm group.

Conversational rhythm

Based on the same set of 1 to 7 syllable groups or feet, conversational rhythm is represented using two different sets of PMPs by the overall, relatively regularly spaced pattern of feet in a clause or sentence.

Discourse features

After PMP's are introduced to learners, they have the chance to practice the structures in short conversations called "two by sixes" which are six turns each practiced in pairs. These conversations give learners the chance to practice pronunciation at a higher discourse level. Discourse features that might arise out of this are final falls (mentioned above in the intonation section), contrastive stress, discourse final lengthening and conversational rhythms that go between speakers and not just within the utterance.

Consonants

A dozen or so consonants that typically carry a high functional load for achieving basic intelligibility are addressed in optional mini-modules.

EHIEP IN THE CLASSROOM AND BEYOND

This section discusses two features of the EHIEP system that have been shown to be consistently effective in a second language/ESL learning context, plus a brief, informal observation as to the positive impact that haptic learning and teaching has had on a non-native instructors.

Classroom Correction

The fourth author has implemented the EHIEP system and haptic classroom correction principles in advanced seven-week long academic preparatory courses at a postsecondary institution. The course syllabus covered common academic components such as interactional strategies, a survey assignment, a debate, a final presentation, as well as vowels, consonants, rhythm, and intonation.

Students were introduced to a new pedagogical movement pattern (PMP), given opportunity to practice it during the week, and then reported back on difficulties, results and insights gained.

The class consisted primarily of Asian and Middle Eastern students, many of whom faced typical challenges in distinguishing – both receptively and productively – between tense and lax vowels. Learning and experiencing the lax and tense vowel PMPs in class provided students not only with a haptic anchor but also with a visual picture of the difference in vowel positioning, established as a visualized matrix resembling the IPA vowel chart, filling the visual field in front of them--a standard feature of HICP work. In terms of correction, the instructor used the vowel PMPs extensively, during regular class work and discussion to give feedback and to correct students' pronunciation of words, particularly new vocabulary whenever necessary. Seeing, doing and experiencing these vowels predictably resulted in frequent "Aha!" responses. Pronunciation diaries included consistent reports of awareness of "corrected" and "to be corrected" vowels in spontaneous speaking as well.

Rhythmic Feet Fight Club

"The Rhythmic Feet Fight Club" was also used in a seven-week long intermediate speaking course. Students were first taught to parse text and to identify focal stress. For this particular process, basic guidelines based on Gilbert's *Clear Speech* (2012) were used. Once students were able to apply these general guidelines to fixed text, they were then taught to hold a tennis or golf ball in one hand, or sometimes even putting on boxing gloves, to attend to focal stress and experience the rhythm of the English language using "boxing-like" movement and touch. (See HICPR (2012) blog for example video.)

Once students were introduced to the Fight Club, the technique was later utilized in a variety of speaking tasks or contexts, not as a planned intervention but as "somatic noticing," implemented when focus on conversational rhythm or perhaps a complex idiomatic phrase was appropriate. Students often reported that this particular protocol helped them feel the rhythm of the language when they speak. In fact, it was observed that if students practiced the Fight Club systematically for approximately 5-10 minutes, about three times a week on their own, their overall fluency and their ability to speak rhythmically often improved noticeably within a few days--occasionally even overnight!

Non-native instructor identity

In a graduate, applied phonology seminar at a postsecondary institution, teachers-in-training enrolled in the course, learned, practiced and demonstrated most of the EHIEP protocols as part of their final course assessment. Learning and experiencing these haptic protocols and PMPs has proven particularly beneficial and empowering for the non-native speakers.

One student from Korea, for example, was very quiet and shy at the beginning of the program. From her perspective, the haptic work contributed substantially to a remarkable transformation that took place in terms of her identity, confidence and oral production. In one of her weekly reflective papers, she explained that both the warm-up and intonation PMPs, which she practiced daily, were especially helpful for developing her confidence and expressiveness. Both the instructors and the classmates' observations confirmed the student's report. She even did an impressive, expressive and well-received presentation at a provincial TESL (BC TEAL) conference.

CONCLUSION

It is, of course, nearly impossible to adequately represent in text form, as we have attempted here, what is obviously a very experiential process. Even simply observing EHIEP instruction passively without "dancing" along with the model and students is only enough to give one a preliminary felt sense of the process. What is being taught in terms of linguistic targets in the haptic-integrated format is not in any sense "new." How it is taught is, however, and especially the focus on embodiment as a balance to contemporary, highly cognitive instruction holds real promise. The particular efficacy of haptic engagement in integrating the senses, well established in general "haptics" research, should prove to be an invaluable resource for pronunciation teaching not only in enabling more effective presentation and correction in general instruction but also in working with the integration of sound change into spontaneous speaking.

ABOUT THE AUTHORS

William Acton, PhD, is Dean of Graduate Studies and Director of the MATESOL program at Trinity Western University, Langley, British Columbia. He formally taught in the MATESOL at the University of Houston, 1979-1991, and the Language Faculty at Nagoya University of Commerce (Japan), 1991-2003. His research, presentations and publications generally relate to the role of movement and embodiment in pronunciation instruction. His recent work since 2005 has centered around the development of the Essential Haptic-integrated English Pronunciation (EHIEP) system and the general application of haptic instructional techniques and technology to language instruction. His current research is available on the HICPR blog (http://hipoeces.blogspot.ca/)

William Acton
Trinity Western University
7600 Glover Road
Langley, BC V2Y 1Y1
Canada
+1-604-888-7511
william.acton@twu.ca
http://hipoeces.blogspot.ca/

Amanda Baker, PhD, is Coordinator of the TESOL program at the University of Wollongong in Australia. Amanda's research interests focus on the dynamic relationships that exist between second language (L2) teachers' knowledge, beliefs and practices, especially in the areas of L2 pronunciation, speaking and listening pedagogy. Classroom research and teacher education, as they relate to L2 oral communication, are primary areas of interest in her research. Some of her recent publications have appeared in TESOL Journal, TESL Canada and The Qualitative Report.

Amanda Baker
Faculty of Education
University of Wollongong
Wollongong, NSW 2522
Australia
+ 61 2 4298 1254
abaker@uow.edu.au

Michael Burri is a PhD student in the Faculty of Education at the University of Wollongong in Australia. He also teaches in the MA TESOL program at Trinity Western University and has taught in a variety of contexts in Japan and Canada. Michael served as the TESOL Speech, Pronunciation, Listening Interest Section Chair (2012-13), and his professional interests include pronunciation pedagogy and research, language teacher development and non-native English-speaking teacher (NNEST) issues.

Michael Burri
University of Wollongong
Wollongong, NSW 2522
Australia
+ 61 2 4200 4537
mikeburri@gmx.net

Brian Teaman is a Professor at Osaka Jogakuin University in Osaka, Japan. He is the assistant director of the Research Institute of International Collaboration and Coexistence there. He is also team-teaching in a course for the MATESOL program at Trinity Western University in Langley, British Columbia. His current research interests are in the phonetics and phonology of second language acquisition, pronunciation pedagogy, computer assisted language learning and MASLE (Machine-Aided Spoken Language Evaluation).

Brian Teaman
Osaka Jogakuin College
2-26-54 Tamatsukuri, Chuo-ku
Osaka, Japan 540-0004
+81-6-6761-6492
teaman@wilmina.ac.jp

REFERENCES

Acton, W. (in press). Accent on OEI. In R. A. Bradshaw, R. A. & A. Cook, (Eds.) *Toward integration: One eye at a time --- Clinical applications*. Vancouver, BC: SightPsych Seminars.

Acton, W. (in preparation). *Essential haptic-integrated English pronunciation teaching system* (haptic video instructional modules, student workbook and audio recordings, instructor manual)

Acton, W. (2012). *Post pronunciation pronunciation*. Paper presented at the 2012 TESOL Convention, Philadelphia, PA.

Acton, W. (2010) . *Full-bodied, multiple modality pronunciation teaching*. Paper presented at the 2010 TESOL Convention, Boston, MA.

Acton, W., Baker, A. & M. Burri. (2009) *Haptic approaches to grammar and intonation instruction.* Paper presented at the 2009 TESOL Convention, Denver, CO.

Acton, W. (1997). Direct speaking instruction (and the mora-bound, focal-stress blues). *The Language Teacher, 21*(9), 7-12. Retrieved from http://jalt-publications.org/tlt/files/97/sep/acton.html

Acton, W. (1994). Directed movement in language instruction, *Journal of Communication and International Studies* (Nagoya University of Commerce) 1:43-52.

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly, 18*(1): 71-85.

Asher, J. (1972). Children's first language as a model for second language learning. *Modern Language Journal, 56*, 133-139.

Baker, A. A. (2012). *Integrating pronunciation into content-based ESL instruction*. Paper presented at the 4th annual Pronunciation in Second Language Learning and Teaching Conference, Vancouver, BC.

Brown, A. (2008). Pronunciation and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 197-207). Cambridge: Cambridge University Press.

Burri, M. & Rauser, K. (2011). *Teachable rhythm techniques for the L2 speaking classroom*. Paper presented at BCTEAL Annual Conference.

Burri, M., & Rauser, K. (2010). *Teaching conversational rhythm in speaking classrooms*, Paper presented at the 2010 TESOL Convention, Boston.

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A reference for teachers of English to speakers of other languages* (2nd ed.). Cambridge: Cambridge University Press.

Chela-Flores, B. (2001). Pronunciation and language learning: An integrative approach. *IRAL – International Review of Applied Linguistics in Language Teaching, 39*, 85-101.

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect, 18*(3), 53-70.

Damasio. A. (2000) *The feeling of what happens: Body and emotion in the making of consciousness*. New York: Mariner.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-397.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1997). Pronunciation instruction for "fossilized" learners: Can it help? *Applied Language Learning, 8*, 217-235.

Gendlin, E. (1996). *Focusing oriented psychotherapy*. New York: Guildord Press.

Gilbert, J. B. (1987). Pronunciation and listening comprehension. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 33-39). Washington, DC: TESOL.

Gilbert, J. B. (2012). *Clear speech* (4th ed.). Cambridge: Cambridge University Press.

Hewings, M. (1995). Tone choices in the English intonation of non-native speakers. *IRAL – International Review of Applied Linguistics in Language Teaching, 33*(3), 251-265.

HICPR. (2012) retrieved November 24, 2012, http://hipoeces.blogspot.ca/

Holme, R. (2012). Cognitive linguistics and the second language classroom. *TESOL Quarterly, 46*(1), 6–29. doi: 10.1002/tesq.5

ICPLA. (2013). International Clinical Phonetics & Linguistics Association, retrieved April 21, 2013, http://www.icpla.org/

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Lessac, A. (1997). *The use and training of the human voice: A bio-dynamic approach to vocal life* (3rd ed.). New York: Drama Book Specialists.

Levis, J. M., & Grant, L. (2003). Integrating pronunciation into ESL/EFL classrooms. *TESOL Journal, 12*(2), 13-19.

Liu, L. (2012). An international graduate student's ESL learning experience beyond the classroom. *TESL Canada Journal, 29*(1), 77-92.

Macdonald, D., Yule, G., & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning, 44*(1), 75-100.

Marks, L. (2002). *Touch: sensuous theory and multisensory media*. Minneapolis: University of Minnesota Press.

McCafferty, S. G. (2004). Space for cognition: gesture and second language learning. *International Journal of Applied Linguistics*, 14(1), 148-165. doi: 10.1111/j.1473-4192.2004.0057m.x

McNerney, S. (2011, Nov. 4). A brief guide to embodied cognition: Why you are not your brain. [Web log] Retrieved from http://blogs.scientificamerican.com/guest-blog/2011/11/04/a-brief-guide-to-embodied-cognition-why-you-are-not-your-brain/

Morley, J. (1991). The pronunciation component in Teaching English to Speakers of Other Languages. *TESOL Quarterly, 25*(3), 481-520.

Munro, M.J., & Derwing, T.M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*(4), 520-531.

Murphy, J. (1991). Oral communication in TESOL: Integrating speaking, listening and pronunciation. *TESOL Quarterly, 25*(1), 51-75.

Naiman, N. (1992). A communicative approach to pronunciation teaching. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 163-171). Oxford: Oxford University Press.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency effects and the emergence of linguistic structure* (pp. 137-157). Amsterdam: John Benjamins.

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /r/ by Japanese learners of English. *Language Learning, 62*(2), 595-683.

Scales, J., Wennerstrom, A., Richard, D., & Wu, S. H. (2006). Language learners' perceptions of accent. *TESOL Quarterly, 40*(4), 715-738.

INTEGRATING FLUENT PRONUNCIATION USE INTO CONTENT-BASED ESL INSTRUCTION: TWO CASE STUDIES

Amanda Baker, University of Wollongong

This paper examines the teaching practices of two ESL teachers who integrated pronunciation instruction into intermediate-level, content/task-based, oral communication courses in an Intensive English program. These practices are analysed according to five categories of pronunciation instruction: Language Awareness; Controlled Practice; Guided Practice, Fluency Development and Free Practice. The fourth category, Fluency Development, is defined based on the work of Nation and Newton (2009), who list, among several criteria, pressure to perform at greater speeds or at least "more smoothly" as central to improving learner fluency. Based on classroom observations, interviews with teachers and student questionnaires, the degree to which the teachers integrated pronunciation according to each of these five categories is investigated. Findings demonstrated that systematic integration of pronunciation instruction into content-based curriculum can be problematic in general and that specific focus on fluency development seems to receive relatively little attention in the classroom. Preliminary recommendations for enhancing systematic pronunciation integration and increasing pronunciation fluency development into curriculum are provided.

The last decade has witnessed increased interest and growth in pronunciation teaching in ESL curriculum due, at least in part, to continued advocacy from not only specialists in this area (Levis & Grant, 2003), but also ESL students (Kang, 2010), for whom intelligible speech is an important goal. Prior to this time, there appears to have been a near 25-year gap where pronunciation pedagogy and focus-on-form in general were de-emphasized. The emergence of Communicative Language Teaching in the late 1970s was to varying degrees responsible for this trend, resulting in the relative exclusion of pronunciation in many language teaching circles (Brown, 2008; Derwing & Munro, 2005; Levis, 2006; Levis & Grant, 2003; Murphy, 1991; Naiman, 1992). Despite this neglect, it has regained ground in language education according to reports that an increasing number of teachers have received training in pronunciation pedagogy (Foote, Holtby, & Derwing, 2011).

During the past two decades, communicative approaches to teaching have evolved as well. Especially in programs focusing on English for Academic Purposes (EAP), communicative approaches such as Task-Based Language Teaching (TBLT) and Content-Based Instruction (CBI) have earned considerable currency. One important question that arises is how pronunciation instruction can be successfully integrated into such programs. As Celce-Murcia, Brinton, Goodwin and Griner (2010) note, "Usually, teachers must balance the needs of their students within a somewhat fixed curriculum. If this is the case, pronunciation is not always explicitly included even in a speaking course, and teachers need to find ways to integrate pronunciation into existing curriculum and textbook materials" (p. 381). From one perspective, it is difficult, if not impossible, to address all the difficulties that language learners experience with their speech; there is simply insufficient time to focus on the wide range of potentially problematic features of English pronunciation encountered by an entire class of L2 learners from a variety of L1 backgrounds in a single course (Munro & Derwing, 2006). In the case of EAP programs that offer courses that separate written and oral communication skills, however, there is greater opportunity to give more attention to sounds and prosody as combined with both general speaking and listening skills. Throughout the last

three decades, teacher educators have strongly advocated an integrated approach to pronunciation teaching in oral communication (OC) curriculum (Celce-Murcia et al., 2010; Chela-Flores, 2001; Gilbert, 1987; Murphy, 1991; Wong, 1987). In fact, several of these specialists have also argued for enhanced incorporation of pronunciation in OC courses, and not simply treating it as an isolated module (Scales, Wennerstrom, Richard, & Wu, 2006). Brown (2008) explains that:

> Many learners, and unfortunately many teachers, treat pronunciation as if it were a separate aspect of language learning. If pronunciation is explicitly handled at all, it is often covered in class slots divorced from the rest of the syllabus. However, given that pronunciation is an indispensable aspect of communicating in speech, and given that successful communication is the basic aim of language learning, pronunciation should be seen as relating to various other communicative aspects of language (p. 203).

To gain a better understanding of how phonology can be integrated in the EAP classroom, an examination of teachers' cognitions (i.e., beliefs and knowledge) and actual teaching practices can provide invaluable insight. To date, several studies have investigated teachers' cognitions, but most have relied on questionnaire (Hismanoglu & Hismanoglu, 2010; Saito, 2011; Saito & van Poeteren, 2012; Sifakis & Sougari, 2005) or interview data alone (Baker, 2011a; Jenkins, 2007; Macdonald, 2002), and none have examined how pronunciation may play an integral role into EAP programs specifically.

RESEARCH QUESTIONS

1. How (and to what degree) do ESL teachers who teach the same intermediate-level OC course integrate pronunciation into content/task-based teaching?

2. What types of pronunciation activities[1] do they use and how frequently do they use them?

3. What do teachers believe about integrating this skill into their courses?

RESEARCH DESIGN

Participants

Two teachers who taught the same OC course in an Intensive English Program in the USA participated in this study[2]. The two teachers, Abby and Ginger (pseudonyms), were both experienced ESL teachers with six and 14 years teaching experience respectively, and both had taught this course at least three times in past semesters. In addition, Abby was bilingual in Portuguese and English, having grown up in Brazil.

Context and Curriculum

---

[1] The terms activities and techniques are used, for the most part, interchangeably throughout this paper. While "activity" generally refers to everything that students may do in the classroom, the term "technique" is a "subordinate" term referring to activities that are "planned and deliberate" that either students or teachers may do (Brown, 2007, p. 180).

[2] The study discussed here was a subcomponent of a larger project that included an examination of the beliefs and practices of five experienced ESL instructors (Baker, 2011b, 2013). However, this paper focuses on the two teachers who both taught the same course to facilitate the analysis of their beliefs and practices.

This intermediate-level EAP course focused on teaching listening, speaking and pronunciation skills using content and task-based instruction with American History serving as the subject matter for the course. The OC course consisted three, 50-minute lessons/week over a 14-week period. The course followed a mainly fixed, structured curriculum based on an in-house study guide plus additional content and activities developed by the individual teachers.

Methods

Three types of data were collected from the teachers: three semi-structured interviews (SSIs) held at approximately the beginning, ¾ point and end of the semester; four classroom observations; and two stimulated recall interviews (SRIs). Two consecutive lessons in the first half of the semester and two consecutive lessons in the second half of the semester were observed, video-recorded and transcribed by the researcher/author. From these video-recordings, the researcher identified segments related to pronunciation instruction. With 48 hours of the second of the two consecutively observed classes, an SRI was conducted, which involved the viewing of the selected pronunciation-oriented segments. During the SRI, the teacher was asked to comment on what she remembered thinking at the time the video recording took place.

Data Analysis

The analysis and coding of pronunciation activities is discussed in detail in Baker (2013) and based on Crookes and Chaudron's (2001) taxonomy of language teaching techniques with each technique being classified as either as controlled, guided (semi-controlled) or free technique. In this paper, these techniques are re-categorized into five broader categories as follows:

1. Language Awareness
2. Controlled Practice
3. Fluency Development
4. Guided Practice
5. Free Practice

The reason for this re-categorization is to focus on the teaching purpose of the technique as it relates to understanding, practice and fluent use of a feature of language, in this case different elements of English pronunciation. Subsumed under *Language Awareness*, any technique that involves the explanation and modelling of pronunciation features, listening discrimination activities, visual recognition activities and actual instructions for other types of pronunciation-related activities is included. Under *Controlled Practice*, repetition drills, including those done accompanied by a specific physical movement or in response to a visual or text-based prompt are included. *Guided Practice* activities can involve information gap activities, referential question-answer activities, preparation work for presentations or dramas, and other semi-structured activities. *Free practice*, in comparison, involves less structured activities such as games, dramas, presentations and discussions. Finally, *Fluency Development* activities are structured or semi-structured activities that focus on helping learners to achieve "automated fluency," which Gatbonton and Segalowitz (2005) define as "the smooth and rapid production of utterances, without undue hesitations and pauses, that results from constant use and repetitive practice" (p. 326). For an activity to be categorized as *Fluency Development*, Nation and Newton's (2009, pp. 152-153) three conditions were used:

1. The activity is meaning-focused.
2. The learners take part in activities where all language is within their previous experience.
3. There is support and encouragement for the learner to perform at a higher than normal rate.

Using a transcription and analysis program called *Transana*, the video footage was coded according to the categories described above. *Transana* then provided a visual display of the timeline of different activity types used in each lesson as well as summarized the duration of each activity (Figure 1).
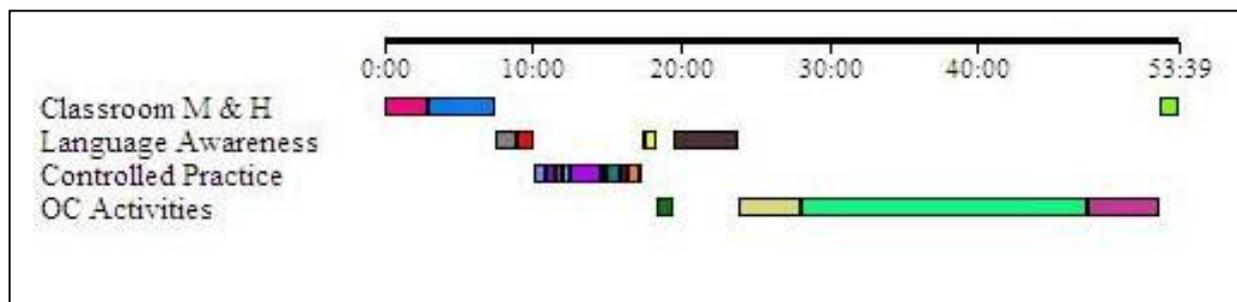
Figure 1. Ginger - Lesson 3 timeline.

Note: Classroom M & H refers to Classroom Management and Housekeeping and pertains to activities such as giving announcements, taking attendance, etc. OC activities refers to any non-pronunciation activities that focused on other general listening and speaking skills, such as taking notes on a lecture or answering questions about a lecture.

FINDINGS

This study revealed how two experienced teachers, to varying extents, integrated pronunciation instruction in their TBLT/CBI oral communication courses. In response to the first research question, pronunciation was an integral component of their classes. This finding is not surprising given that the course curriculum requires pronunciation to be taught in each module of the course. However, although the two instructors teach the same OC course, they integrate phonological features in different ways and to differing degrees. Table 1 shows the percentage of time dedicated to developing pronunciation in the two classes, based on the observations and interviews.

Table 1

Observations and Reports of Integration of Pronunciation Instruction

	Abby	**Ginger**
Observation of 4 lessons	69.7% (160 min 30 s)	17.4% (37 min 42 s)
Teacher's Perception[3]	60-70%	20% in-class (higher % outside class)

---

[3]During the second SSI, teachers were asked to respond to the following question: "What percentage of time do you think you spend on pronunciation in your class?"

The two data sources on the average time spent on pronunciation varied across the two classes, showing that Abby's class devoted considerably more time overall. These results indicate that pronunciation is integrated with other OC skills, but has a dominant role in her lessons. In her first interview, Abby explained this integration:

> We use American government as the basis, so I try to do something with the content. Half the class content and half the class with pronunciation. There's only 50 minutes so sometimes it really doesn't work...but I try to stick in some kind of pronunciation aspect even like syllables. Either we're practicing syllables or we're counting syllables. Today word stress started - the rules for word stress; Friday [we will] continue that. And then part of the class next time on Monday, [we will] continue finishing up the rules. Today I also did the lecture, so Monday I repeat the lecture. That's just part of [this course in the] IEP. And so all the time I try to do something pronunciation and then something content based.

However, Abby also emphasized that she intertwined pronunciation with the course content whenever possible, and that "the content gives it a sort of continuum, something to link it to."

> I think [pronunciation] should be just pretty much all the time….How I say the question or how the students are saying the question and reviewing that pronunciation or words within their answer, words within the question, word stress, endings if we are doing -ed endings, how to say that. Even in the PowerPoint, if there's the word "representative" up there...so where's stress...just very quickly while I'm doing the lecture...where's the stressed syllable? I guess any chance you get reinforcing what the pronunciation feature is for that unit or lesson or chapter.

She further noted that:

> I have to say that with the content, [CBI] does make it easier to teach just simply because you have something to … attach all these skills to […] The content gives it a sort of continuum, something to link [pronunciation] to.

In comparison with Abby, Ginger spent less time teaching pronunciation; however, her lessons still reflected an integrated approach to teaching OC skills. Similar to Abby, she strove to merge pronunciation with the subject matter of the course, frequently focusing on key words from lectures on American history and having students practice their pronunciation. She explained that:

> I have them repeat and they do drill and practice stuff...and then I do checks with them. I just call on them and they have to read the words and then if it's wrong, then I'll say something, and then I'll repeat it, and then I'll have the whole class repeat and not just that one student. I feel like it's my job in pronunciation time to highlight discrepancies, but otherwise I try to save that kind of thing for WIMBA.

Her reference to WIMBA refers to an online program that Ginger used to give students feedback on their pronunciation. The learners recorded key words and sentences using the audio record feature, and later she listened to the recording and then provided feedback on their pronunciation. She mentioned that she spent considerable time giving students feedback on their pronunciation, which may account for why the students reported spending a great deal of time on this skills in the course (Table 1).

In response to the second research question, the teachers used Language Awareness activities, Controlled Practice, Free Practice and Guided activities according to observational and interview data. Table 2 provides the results of the frequency of the activities based on observational data alone. It is important to note here that although the teachers' use of guided and free activities is almost nonexistent in the observational data, the use of these types of activities does surface in the interview data, albeit to a limited extent (see Baker, 2013, for an expanded discussion).

Table 2

Frequency of Activities in Observed Classes

Activity Type	Abby		Ginger	
Class. M & H	8.2%	20 min	10.6%	23 min
L. Awareness	**49.4%**	**114 min**	**8.6%**	**19 min**
Controlled P.	**5.5%**	**13 min**	**6.8%**	**14.5 min**
Guided P.	**8.5%**	**19 min**	**0**	
Fluency D.	**0.5%**	**1 min**	**0**	
Free P.	**0**		**0**	
OC Activities	27.8%	64 min	73.9	160 min
Total Time		230 min		216.5 min

Based on the combined data from the interviews and the observations, the one type of activity that appears to be virtually missing from the content-based ESL classes is Fluency Development. As described by Nation and Newton (2009, pp. 152-153), fluency development activities are "meaning-focused," include only language from "within their previous experience" and require students to "perform at a higher than normal rate." Based on this definition, activites specifically devoted to fluency development rarely, if ever, surfaced in the data, thus suggesting a possible missing link in the students' development of fluent language use.

In response to the third research question, one concern raised by the teachers during the interviews related to a struggle for balance. Both Abby and Ginger highlighted the difficulty of providing sufficient time to course content focus and to pronunciation. Ginger said:

> I feel like I try to incorporate as much pronunciation as possible while still getting through the content of the course.... It's supposed to be a listening, speaking, pronunciation type course and I try to find balance among all of that.

Similarly, Abby explained that "The problem with this class is that you always have to juggle content time and pronunciation time."

Another concern that was discussed in depth by Abby was how best to integrate pronunciation into the course. Over several semesters, Abby experimented with different ways to address pronunciation instruction. For some units, she tried to disperse pronunciation instruction throughout the entire unit. In other units, she used a sandwich approach, alternating content and pronunciation, before finally reviewing the content again at the end of a unit. At the end of the current study, Abby concluded that she preferred the sandwich approach, feeling that it seemed more effective. She said:

> In this last unit, I introduced all the content first, and then I focused on the pronunciation. It seemed to work. I kind of liked that. So I'll try it again and see if I really do like it that way. Present all the content first and then the pronunciation, and then just kind of review content at the end.

That said, pronunciation instruction still occurred throughout the units, as noted by Abby earlier in this paper. Frequently she would take every opportunity to focus briefly on pronunciation even when teaching American History.

DISCUSSION AND CONCLUSION

Overall, the observations and the teachers' perspectives indicate that in principle the classes follow an inclusive curriculum targeting not just general speaking and listening skills, but pronunciation skills as well. The position of pronunciation as integrated with other OC skills seemed to have especially high priority in the classes taught by Abby, the teacher who is the bilingual Portuguese-English speaker. Whether having a large or small focus in the teachers' courses, however, it is apparent that they both highlighted the importance of this skill. Their belief is likely shared by numerous teachers around the globe who work in EAPs, but particularly those EAPs that value focus on form, recognizing the critical role that prosody and sounds can play in intelligible speech.

The results also revealed a lack of activities specifically devoted to pronunciation-oriented fluency development within the TBLT/CBI classes. While pronunciation instruction and practice was integral to the course and to overall language development, the dearth of fluency-development activities in particular may likely inhibit learners' development of "automated fluency" (Gatbonton & Segalowitz, 2005) of intelligible features of pronunciation. If it is the case that fluency may be negatively affected, or at least not advanced, it may be beneficial for teachers with similar OC courses to add activities that specifically focus on improving the fluency of intelligible pronunciation in learner speech. There are any number of possible activities that can aid in this area, including haptic (movement + touch) techniques such as the *Rhythmic Feet Fight Club* (Acton, Baker, Burri, & Teaman, forthcoming) or other general fluency activities such as the 4/3/2 technique (see Arevart & Nation, 1991) – as long as the learners are directed to focus on intelligible pronunciation (see Nation & Newton, 2009, for additional suggestions).

Finally, both teachers expressed a concern with the difficulty with maintaining balance between content and language development. Uniting pronunciation instruction with speaking and listening skill development, while at the same time teaching subject content matter, was considered a challenge. Nevertheless, they asserted that in their experience the content can provide a communicative anchor for pronunciation instruction. As research has yet to be conducted on how to effectively and efficiently integrate pronunciation instruction into CBI

or TBLT, the teachers' expressed need highlights another important "opportunity" for class-based research. Quasi-experimental research has explored the teaching of pronunciation in ESL classrooms (Couper, 2003), but comparisons are needed to explore different paths to fluency in the integrated L2 classroom.

To conclude, the results of this research indicate that pronunciation instruction has a definite role in the teaching of OC skills in CBI/TBLT, at least in the classrooms of these instructors. Future research needs to address not only how to integrate pronunciation effectively in CBI curriculum, but also how to incorporate activities that specifically target the development of fluent, intelligible pronunciation. For this target to be fully realized, however, particular attention must be given to the world beyond the classroom, that is, the robust development of pronunciation skills that are able to withstand the rigors of the local Starbucks (or elsewhere).

ABOUT THE AUTHOR

Amanda Baker, PhD, is Coordinator of the TESOL program at the University of Wollongong in Australia. Amanda's research interests focus on the dynamic relationships that exist between second language (L2) teachers' knowledge, beliefs and practices, especially in the areas of L2 pronunciation, speaking and listening pedagogy. Classroom research and teacher education, as they relate to L2 oral communication, are primary areas of interest in her research. Some of her recent publications have appeared in TESOL Quarterly, TESOL Journal, and TESL Canada.

REFERENCES

Acton, W., Baker, A. A., Burri, M., & Teaman, B. (forthcoming). Preliminaries to haptic-integrated pronunciation instruction. In J. M. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference, Aug. 2012.* Vancouver, BC.

Arevart, S., & Nation, P. (1991). Fluency improvement in a second language. *RELC Journal, 22*(1), 84-94. doi: 10.1177/003368829102200106

Baker, A. A. (2011a). Discourse prosody and teachers' stated beliefs and practices. *TESOL Journal, 2*(3), 263-292. doi: 10.5054/tj.2011.259955

Baker, A. A. (2011b). ESL teachers and pronunciation pedagogy: Exploring the development of teachers' cognitions and classroom practices. In J. M. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference, Sept. 2010* (pp. 82-94). Ames, IA: Iowa State University.

Baker, A. A. (2013). Exploring teachers' knowledge of L2 pronunciation techniques: Teacher cognitions, observed classroom practices and student perceptions. *TESOL Quarterly, Advance online publication*. doi: 10.1002/tesq.99.

Brown, A. (2008). Pronunciation and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 197-207). Cambridge: Cambridge University Press.

Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy* (3rd ed.). White Plains, NY: Pearson Education.

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A reference for teachers of English to speakers of other languages* (2nd ed.). Cambridge: Cambridge University Press.

Chela-Flores, B. (2001). Pronunciation and language learning: An integrative approach. *IRAL, 39*, 85-101.

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect, 18*(3), 53-70.

Crookes, G., & Chaudron, C. (2001). Guidelines for language classroom instruction. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (3rd ed., pp. 29-42). Boston, MA: Heinle and Heinle.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-397.

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29*(1), 1-22.

Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *The Canadian Modern Language Review, 61*(3), 325-353. doi: 10.1353/cml.2005.0016

Gilbert, J. B. (1987). Pronunciation and listening comprehension. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 33-39). Washington, DC: TESOL.

Hismanoglu, M., & Hismanoglu, S. (2010). Language teachers' preferences of pronunciation teaching techniques: Traditional or modern? *Procedia: Social and Behavioural Sciences, 2*(2), 983-989. doi: 10.1016/j.sbspro.2010.03.138

Jenkins, J. (2007). *English as a lingua franca: Attitude and identity* Oxford: Oxford University Press.

Kang, O. (2010). ESL learners' attitudes toward pronunciation instruction. In J. M. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (pp. 105-118). Ames, IA: Iowa State University.

Levis, J. M. (2006). Pronunciation and the assessment of spoken language. In R. Hughes (Ed.), *Spoken English, TESOL and applied linguistics: Challenges for theory and practice* (pp. 245-270). New York: Palgrave MacMillan.

Levis, J. M., & Grant, L. (2003). Integrating pronunciation into ESL/EFL classrooms. *TESOL Journal, 12*(2), 13-19.

Macdonald, S. (2002). Pronunciation views and practices of reluctant teachers. *Prospect, 17*(3), 3-18.

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*, 520-531.

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning, 58*(3), 479-502.

Murphy, J. (1991). Oral communication in TESOL: Integrating speaking, listening and pronunciation. *TESOL Quarterly, 25*(1), 51-75.

Naiman, N. (1992). A communicative approach to pronunciation teaching. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 163-171). Oxford: Oxford University Press.

Nation, P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge.

Saito, K. (2011). Identifying problematic segmental features to acquire comprehensible pronunciation in EFL settings: The case of Japanese learners of English. *RELC Journal, 42*(3), 363-378. doi: 10.1177/0033688211420275

Saito, K., & van Poeteren, K. (2012). Pronunciation-specific adjustment strategies for intelligibility in L2 teacher talk: results and implications of a questionnaire study. *Language Awareness, 21*(4), 1-17. doi: 10.1080/09658416.2011.643891

Scales, J., Wennerstrom, A., Richard, D., & Wu, S. H. (2006). Language learners' perceptions of accent. *TESOL Quarterly, 40*(4), 715-738.

Sifakis, N. C., & Sougari, A.-M. (2005). Pronunciation issues and EIL pedagogy in the periphery: A survey of Greek state school teachers' beliefs. *TESOL Quarterly, 39*(3), 467-488.

Wong, R. (1987). Learner variables and prepronunciation considerations. In J. Morley (Ed.), *Current perspectives on pronunciation: Practices anchored in theory* (pp. 17-28). Washington, DC: TESOL.

PRONUNCIATION FROM THE PERSPECTIVE OF PRE-SERVICE EFL TEACHERS: AN ANALYSIS OF INTERNSHIP REPORTS

Larissa Buss, Concordia University

This pilot study analyzed internship reports written by seven pre-service English as a foreign language (EFL) teachers in Brazil with the purpose of investigating the student teachers' knowledge, beliefs, and practices regarding pronunciation. Data analysis involved reading through the reports and coding the excerpts in which pronunciation was addressed. These codes were then collapsed into four broader themes (Creswell, 2012): *identifying problems*, *explaining problems*, *dealing with problems,* and *expressing beliefs*. The findings suggest that the student teachers had unclear knowledge of the phonologies of English and Brazilian Portuguese (BP) and that they were not fully aware of the reasons behind common pronunciation errors made by Brazilian EFL learners. Overall, pronunciation was taught mainly through modelling and repetition, and greater attention was given to segments or individual words. These results differ from previous research findings in ESL settings, where teachers reported using more varied techniques and teaching a combination of segmentals and suprasegmentals (Breitkreutz, Derwing, & Rossiter, 2001; Burgess & Spencer, 2000; Foote, Holtby, & Derwing, 2011). They also suggest that the pre-service teachers in Brazil were not adequately prepared to teach pronunciation, which might point to a flaw in their university education.

INTRODUCTION

It is well-known that research and practice do not always walk hand-in-hand. In education, this discrepancy can be quite large, since many practitioners are unaware of the latest research findings or do not know how to apply them in their classrooms (Borg, 2009). Studies investigating teachers' knowledge, beliefs, and practices can provide valuable information about the reality of education in a given area and may help to identify problems that require urgent attention. In the domain of L2 pronunciation, however, research on teachers' cognitions and classroom practices is scarce (Baker & Murphy, 2011). Thus, it is still unclear how recently renewed interest in pronunciation has influenced teachers worldwide, especially in EFL settings. Fortunately, research aimed at answering this question is growing, and this study provides a contribution.

PRONUNCIATION INSTRUCTION AND TEACHERS' BELIEFS

For over a decade, researchers have identified a need for more teacher training opportunities in pronunciation, as many instructors do not feel completely comfortable teaching it in the English as a second language (ESL) classroom (Baker, 2011; Foote et al., 2011; Macdonald, 2002). Yet some studies suggest that ESL teachers' beliefs and practices are generally in line with current pronunciation research, at least in North America and in the UK. For example, ESL instructors have been found to value pronunciation teaching at all proficiency levels and to regard intelligibility as a more appropriate goal than accent reduction (Breitkreutz et al., 2001; Foote et al., 2011). Regarding teaching practices, Baker (2011) concluded that research on pronunciation strongly influenced the teaching of five of the six instructors in Canada and the USA. These teachers tended to prioritize prosodic features, according to what they had learned in their

graduate education programs. The surveys conducted by Breitkreutz et al. (2001) and Foote et al. (2011) found that Canadian teachers reported teaching both segmental and suprasegmental features and using pronunciation textbooks in their classes. In Foote et al. (2011), the instructors mentioned using a variety of techniques and resources that they considered effective for improving learners' pronunciation, including games with minimal pairs, repetition, mirrors, visual aids, and tactile reinforcement. Although most of these activities worked primarily on segmentals, working on intonation was considered important, and several suprasegmental features were mentioned as problematic for ESL learners (Foote et al., 2011). Similarly, British teachers questioned by Burgess and Spencer (2000) reported making use of a number of strategies to deal with pronunciation problems, including the phonemic alphabet, drills, chanting, drama, and role-play. Suprasegmental features like intonation and stress were seen as highly important, as well as certain segmental aspects (e.g., voicing, schwa, clustering, and linkage phenomena) and their relationship to suprasegmentals. Stress, rhythm, and intonation were regarded as main areas of difficulty for students.

To date, little is known about pronunciation teaching in the numerous countries where English is taught as a foreign language (i.e., countries where English is not a primary language). In typical EFL contexts, learners share the same L1 and have fewer opportunities to speak the target language in real-world situations or to interact with native speakers. EFL teachers usually do not have the challenge of dealing with multilingual groups, but they certainly have other difficulties to face. In EFL settings like Brazil, the vast majority of teachers are non-native English speakers. One could speculate that the teachers' own difficulties with some aspects of oral English could make them uncomfortable in teaching pronunciation. Another important aspect mentioned by Derwing (2008) is that students who share the same L1 can sometimes be more intelligible to one another when speaking the foreign language. Thus, they might need less effort to make themselves understood and could create a distorted idea of how clearly they speak. Even teachers who share the students' L1 may be subject to this false impression. These and other particularities of EFL settings might have implications for teachers' views and practices concerning pronunciation. However, most research conducted with non-native instructors has focused almost exclusively on beliefs, particularly on their attitudes toward native models and accented speech. An interesting finding in such studies is that non-native teachers appear to have contradictory opinions. While they admire native pronunciation as an ideal, they also recognize the central role of intelligibility (Sifakis & Sougari, 2005; Timmis, 2002) and they understand that accent is an important marker of L1 identity (Jenkins, 2005).

A recent project that has examined EFL teachers' attitudes and classroom practices is the English Pronunciation Teaching in Europe Survey (EPTiES). According to preliminary findings published by Henderson et al. (2012), teachers from seven European countries attributed high importance to pronunciation and rated their own pronunciation favourably. Results of the EPTiES for Finland revealed that most instructors from that country taught learners to recognize phonetic symbols and used at least some ear training as part of their pronunciation teaching (Tergujeff, 2012a). A case study conducted by Tergujeff (2012b) found that Finnish teachers emphasized segmental features and adopted traditional pronunciation teaching practices. The main activities used by the four instructors observed were listen and repeat tasks, and teachers often corrected their learners or pointed out the correct pronunciation of words. Read-aloud tasks and phonemic script were sometimes used by two of the teachers, but other types of activities were rare.

Similar research in other EFL settings is clearly warranted. The current study partially addresses this gap by investigating the cognition and practices of pre-service EFL teachers in Brazil. The main research questions are the following:

1. What pronunciation teaching practices do the student teachers report using?

2. What beliefs regarding pronunciation teaching and learning do they express?

3. How much knowledge do they demonstrate of English pronunciation, pronunciation instruction, and the main difficulties experienced by Brazilian EFL learners?

While previous studies have typically used surveys, questionnaires, and interviews to elicit responses from teachers, this study analyzed internship reports. The rationale behind this choice is that internship reports can provide contextualized information about student teachers' practices and cognition. In addition, the participants did not know in advance that the researcher was interested in pronunciation, so they were not necessarily prompted to include it in their practices and reflections.

METHOD

Participants and Materials

The data used in the study was made available to the researcher by a federal university in the South of Brazil. It consisted of seven internship reports written by last-year undergraduates studying in a Portuguese and EFL teacher education program. The internship supervisor, a tenured faculty member at the university, provided information about the undergraduate program and the teaching internships. The researcher was also given access to the program flowchart and to the course outline with the guidelines given to the students for the write-up of their teaching internship reports.

The undergraduate program is four years long and provides students with a broad background in Portuguese, English, language teaching, and literature. A significant portion of the course load consists of EFL classes, as many students enter the program with low proficiency in English. These classes and the courses in EFL teaching offered in the program are largely based on the principles of communicative language teaching (CLT). The student teachers in this study had taken a sixty-hour course in phonetics and phonology in their first year. They also received some instruction and guidance on how to teach oral skills within more general courses, such as Applied Linguistics, but pronunciation teaching was not covered specifically.

The teaching internship was a mandatory component of the program. It consisted of one semester of supervised EFL teaching covering the four language skills (reading, writing, listening, and speaking). One participant taught a group of teenagers aged 16 to 17 at a local private high school, while the others taught mostly young adults taking part in non-credit courses offered by the university to the community. The average group size was 15 learners (range: 9-20), all native speakers of Brazilian Portuguese (BP) and beginners in English. Although this was their first official teaching internship in the program, all of the student teachers had previous English teaching experience (e.g., at private language schools).

The internship reports covered 18 to 20 hours (i.e., approximately one month) of teaching in the middle of the course. That is, although the student teachers taught an entire semester, they focused their reports on this one month of teaching, a period that had been assigned by the internship supervisor for evaluation purposes. The reports included lesson plans and descriptions

of each class, as well as critical reflections. The evaluation criteria regarding the content of the reports were the following: (1) critical reflection on teaching practices with reference to relevant theory and literature, (2) clarity of assessment criteria, and (3) reflection and discussion about learners' main language "errors/problems." The student teachers were told not only to describe their practices and their learners' performance, but to justify the former and discuss possible reasons for the latter from an informed perspective. The reports were written in BP and their average length was 47 pages, excluding appendices (e.g., photocopied activities from books, learners' homework, marks, etc.).

Data Analysis

Data analysis followed the inductive procedure described in Creswell (2012). After an initial exploratory analysis of the entire reports, the researcher located all the text segments about pronunciation and coded them for emerging topics. These codes were then collapsed into four broader themes: *identifying problems*, *explaining problems*, *dealing with problems,* and *expressing beliefs*, which are explained and illustrated in the next section. In order to ensure confidentiality, each teacher is referred to with a number (from T1 to T7). Direct citations have been translated into English by the author.

RESULTS

The preliminary exploratory analysis revealed that all of the student teachers mentioned pronunciation in their internship reports. However, the number of references to the topic varied considerably, ranging from only two short passages (T4, T5) to as many as 21 passages (T7). Together, the comments related to pronunciation totalled 3,652 words. The main findings of the analysis are presented by theme below.

Identifying Problems

All of the teachers but one (T4) reported that their students experienced pronunciation difficulties at some point during the course. Yet numerous times they did not explain what these problems were, only mentioning generically that students had "many difficulties with pronunciation" (T3, T6, T7) or that "some/many pronunciation problems" were observed during specific oral activities (T1, T2, T3). Two teachers referred to "serious" (T1, T7) and "normal" (T7) pronunciation problems without indicating what these terms meant (for example, whether they thought an error was serious because it could hinder communication or whether another problem was considered normal because of specific features in the students' L1).

On the other hand, some problems were commented on more specifically. Three things were reported as particularly challenging: the *th* sound /θ/ in ordinal numbers (T1, T3), /r/ in word-initial position (T3, T6), and the *-ed* inflectional endings for regular verbs in the past tense (T1). Teacher 1 also mentioned that her students pronounced *were* like *where*, and Teacher 2 listed 11 words that her learners mispronounced with phonetic transcriptions of how they pronounced them. The epenthesis of [i] was observed at the end of some of these words (e.g., *gave* being pronounced as [geɪvi]), and Teacher 2 specifically mentioned this feature as a common problem in her learners' speech. Teacher 6 noticed the same phenomenon when her students pronounced the word *band*, for example. Only Teacher 7 identified fluency and intonation as troublesome aspects. In particular, he noticed that some students' reading was monotonous because they did not use proper intonation.

In many cases, however, the teachers referred to problematic words without identifying where exactly the problems lay. For example, they said "there was difficulty with the pronunciation of *you're welcome*" (T6), "I noticed difficulty… in the pronunciation of the parts of the body, such as *thumb*, *ankle* and *knee*, among others" (T5), and "some problems appeared… regarding the pronunciation of health problems" (T1). It is unclear whether the teachers were aware of the difficulties and just did not report them or if they only had a general impression that something was wrong in their learners' pronunciation.

Explaining Problems

Even when the student teachers were able to pinpoint pronunciation difficulties, they did not usually attempt to explain their causes. Unfortunately, the explanations that could be found were generally vague or flawed. On three occasions, the teachers simply attributed the errors to transfer or to "the influence of Portuguese" (T2, T5), without further elucidation. Another common claim was that students' difficulties were caused by the differences between the sounds of English and BP (T2, T3, T6). In reality, however, other factors could account for the errors, including different syllable structure and opaque orthography.

Some student teachers seemed to overgeneralize problems and to mix up the concepts of sounds and letters in their explanations. For example, Teacher 2 wrote that "many students tend to put an [ɪ] sound at the end of words ending in consonants or in the vowel [e], because of the influence of Portuguese." Actually, epenthesis does not occur in all words ending in consonants, as BP admits some consonants in word-final positions: /l/, /r/, /N/, and /s/. Thus, Brazilian EFL learners will often add an [i] at the end of words like *dog* and *map* (i.e., they will say [dɑgi] and [mæpi]), but not after *miss* or *car*, for example. Moreover, epenthesis would never occur after vowel sounds such as [e]. It is likely that the teacher was actually referring to words that end in a silent letter "e" such as *time* and *because*. In these cases, orthographic interference causes learners to pronounce the final vowel. Confusion is also observed in the following comment by Teacher 3:

> The other [learner] has more specific problems, like sounds that we do not use in Portuguese, such as the sound of 'r' at the beginning of words, the sound of 'th', and the sound of 's' with the sound of 'j', as in the word *usually*, which are characteristic of the English language.

The retroflex /r/ is not present at all in most dialects of BP. The reason why its pronunciation can be especially tricky in word-initial position is because BP speakers produce a fricative r-sound that is similar to /h/ for words beginning in "r". Thus, EFL beginners from Brazil might pronounce the words *rat* and *hat* identically. As for words like *usually*, the main difficulty is that the spelling will induce learners to pronounce /z/ instead of /ʒ/. (In BP, intervocalic "s" is pronounced as /z/.) Nonetheless, /ʒ/ cannot be considered characteristic of the English language, as it is quite common in BP for words written with "j" or "g" followed by "e" and "i". Although Teacher 3 did not express clear awareness of the influence of orthography in the comment above, she did mention that another learner read English words "exactly as they are written, as if she was reading in Portuguese." A similar observation was made by Teacher 6, who noticed that her learners tended to read the word *favourite* "as it is written, pronouncing an [i] at the end." In this word, both the spelling and the final consonant sound may cause the production of epenthesis by Brazilian learners.

Dealing with Problems

When the teachers talked about their instructional practices, very generic descriptions were often used. Rather than describing specific strategies or activities, they would frequently say, for instance, that they did "thorough/intense work on pronunciation" (T1), "focused on pronunciation" (T3), or "helped the students with pronunciation" (T6). Whenever an activity was mentioned, it usually involved listening and repeating the target structure. Teacher 1, for example, said she had to "work arduously on repetition" in order to improve students' pronunciation of certain words. Teacher 3 mentioned having students repeat the *th* many times until most of them got it right. According to Teacher 7, students practiced intonation and stress by imitating a recording and repeating sentences many times. Besides these, several other references to repetition were found in the internship reports. In total, there were 14 mentions of repetition in eight text segments and four reports.

Another very common practice, mentioned eight times and by five teachers, was to have learners read texts and dialogs aloud. During these tasks, the teacher would often monitor and correct students' pronunciation. Teacher 4, for instance, asked her group to read a conversation in pairs and "pay attention to pronunciation and intonation." Then, she walked around the classroom correcting the problems she observed in each pair. It is possible that the teacher did this by simply providing the students with the target-like pronunciation, but this was not clearly described in the report. Read-aloud activities were also used by at least two teachers (T2 and T3) to formally assess learners' oral skills. The practice of monitoring and correcting students' pronunciation was reported a total of nine times by four teachers.

When evaluating her teaching experience, Teacher 1 indicated that repetition exercises were effective at improving her students' pronunciation. She stated that "the great improvement students showed was in terms of pronunciation. Most of the group presented serious problems… I had to work arduously on repetition. However, this work proved to be effective, since there was clear improvement in pronunciation." Conversely, other teachers found that the same strategy did not work for all of the students. Teacher 3, for instance, mentioned that some pronunciation problems persisted regardless of the numerous repetition exercises she did: "Some students had difficulties pronouncing the *th* ending and I made them repeat it many times until they got it right, but two of them did not." Later she reflected: "It is interesting to notice that even with so many exercises of audio and repetition, besides the conversation exercises, these students have so many problems in their speech." Teacher 3 was aware that her teaching was not effective for all the students, but perhaps she did not know what else to do to help them. Similarly, Teacher 7 said that his group generally improved, but some learners could not keep up: "It is possible to see some progress in some students regarding pronunciation and intonation, but in general most of them still need more practice… Unfortunately, some students could not maintain the oral standard achieved by most of the students."

Expressing Beliefs

When looking into the excerpts in which the pre-service teachers expressed their thoughts and opinions, no general tendencies were identified. Still, some interesting topics emerged, including reflections about pronunciation models and teaching objectives. Teacher 1 was particularly concerned about her role as a model for learners:

> I can say that the work done with the students helped me to reflect on my own pronunciation. I became more aware of the mistakes I made and that I can never be too careful, as a large portion of the input students receive comes from the teacher.

She reported having difficulties with the pronunciation of *th* herself and that she had to practice at home before going to class in order to pronounce it "as correctly as possible". Teacher 1 also mentioned having spent an entire class mispronouncing the word *determiners*. When she found out her mistake, she realized that she needed to pay more attention to her own pronunciation.

Teacher 7 emphasized the importance of a native model. He said that all repetition activities were preceded by a recording so that "the students could have the exact perception of how a native speaker of the target language expresses himself, taking into account his pronunciation, accent, and intonation." He also justified his use of drilling activities by explaining that "it is through them that the students will have a model to improve their pronunciation." On the other hand, somewhat paradoxically, Teacher 7 reported telling his learners that English could have many accents and that making oneself understood was the main objective of oral communication. He told them that communication would be achieved as long as the message was conveyed "clearly and with good pronunciation." Teacher 6 expressed a similar idea about the objective of oral practice. She mentioned that the classroom activities aimed to "develop the intelligibility of the students when they communicate" and that "a perfect pronunciation should not be demanded" from the learners, since they have limited English exposure and language acquisition is a slow process. Nonetheless, similarly to Teacher 7, she seemed to hold the idea that pronunciation practice typically involved repetition and that it was disconnected from more communicative tasks: "Although we try to provide moments of authentic interaction, that is, when they can communicate based on their own context, this does not mean we cannot do repetition exercises, whose focus would be the practice of pronunciation".

Teacher 7 and Teacher 3 expressed some disbelief in the effectiveness of pronunciation teaching. Although Teacher 7 valued pronunciation, he did not seem to think that instruction in the classroom was enough for his learners. He claimed, on three occasions, that the learners had so many difficulties because they did not practice oral skills outside the classroom. Teacher 3 seemed somewhat pessimistic about the effects of pronunciation teaching to large groups of learners. She argued three times that pronunciation could not be taught adequately in these contexts, because it is impossible to "hear everyone and everything." She believed that significant improvement was only achieved when the teacher had "enough time to work on students' pronunciation errors one by one," despite the fact that her learners shared the same L1.

DISCUSSION AND CONCLUSION

The data revealed four main themes within the reports. Regarding the first, pre-service teachers often referred to "pronunciation problems" or cited vocabulary that their students had difficulties pronouncing, but did not explain what the exact problems were. This might indicate that sometimes the teachers were unable to articulate the problems they perceived in their students' pronunciation. Furthermore, the great majority did not mention prosodic features like stress, rhythm, and intonation as problematic, which differs from previous research findings obtained in ESL settings (e.g., Burgess & Spencer, 2000; Foote et al., 2011). In fact, certain segments and especially the pronunciation of individual words were the most frequently mentioned problems. It is unlikely that their learners did not struggle with pronunciation beyond the word level. It

might be that the student teachers were not sensitive to these problems because had not been trained to identify and work on them in the classroom.

As for the second theme, it was found that the teachers did not adequately account for the pronunciation problems reported. More often than not, they presented no explanations or generally attributed these problems to the influence of the L1 or to the differences between the sounds of the two languages. In the more detailed accounts, some misconceptions were identified, which suggests that the teachers had little knowledge of the phonologies of BP and English and of the reasons behind very common mistakes made by Brazilian EFL learners.

Unfortunately, some student teachers did not write much about their teaching practices. In several passages, they mentioned having worked on pronunciation without describing specific approaches or exercises. The activities that were described mainly involved listening to a model – either the teacher or a recording – and repeating a target word or structure many times. Asking learners to read aloud and correcting their pronunciation were also frequently reported practices. None of the teachers mentioned using more explicit techniques to raise students' awareness of pronunciation, such as articulatory descriptions, phonetic symbols, or visual aids, not even when repetition was ineffective. It appears that these instructors adopted an *intuitive-imitative approach* rather than an *analytic-linguistic approach*, which would have been more complete (for a description of these approaches, see Celce-Murcia, Brinton, Goodwin, & Griner, 2010). Thus, their practices differ from those reported by ESL teachers in Burgess & Spencer (2000) and Foote et al. (2011), which included an array of techniques to teach pronunciation explicitly. In their pronunciation teaching, the Brazilian student teachers more closely resembled the Finnish EFL instructors in Tergujeff (2012b), who emphasized segments and made use of imitation, teacher correction and, to a lesser extent, reading aloud.

Two teachers were concerned about the need for providing students with accurate pronunciation models. At the same time, the rather contradictory notion that intelligibility is a more appropriate goal for learners also came up. This idea might be a reflection of the communicative framework in which the participants were trained. Previous studies involving non-native English instructors have also identified a conflict between their attachment to native models and their awareness of what is realistic or relevant for learners (Sifakis & Sougari, 2005; Timmis, 2002). Interestingly, the same teachers who mentioned the importance of intelligibility associated pronunciation instruction with repetition tasks, apparently seeing it as something separate from communicative practice.

In conclusion, it is possible to say that the pre-service EFL teachers generally attended to pronunciation and possibly understood the importance and purpose of its teaching. Nonetheless, they did not appear to be fully prepared for this task for two main reasons. First, most of the time they could not appropriately diagnose and explain very common pronunciation errors made by Brazilian EFL learners. Second, their teaching seemed to be restricted to an intuitive-imitative approach, with no use of more explicit, awareness-raising activities. These results suggest a disparity between previous ESL teaching practices reported and the EFL context in this study. In addition, they may point to a need for more specific training in pronunciation instruction at the participants' university.

Such results should be interpreted within the limitations of the present study, which include the small sample size and the fact that internship reports are not necessarily thorough accounts. It is possible that the student teachers forgot to describe some their practices or omitted information

that could be relevant to this study. Another limitation is that their reported practices and cognition concerning other areas of language learning (e.g., English grammar) were not analyzed for comparison. It may be that the student teachers had difficulties in other areas as well, not only pronunciation. Still, it is believed that what they did and did not report provided some indications of their beliefs, knowledge, and practices related to pronunciation. Future research addressing the limitations of this study is called for. In particular, as Baker and Murphy (2011) observed, there is a need for studies on teacher cognition that include a classroom observation component. This would allow for the collection of more detailed and reliable data about teachers' classroom practices.

ABOUT THE AUTHOR

Larissa Buss is a first-year PhD student in Education at Concordia University, in Montreal. She is specializing in Applied Linguistics and her research interests include EFL, teacher education, pronunciation instruction, and second language phonetics and phonology. She holds a full CAPES scholarship for her doctoral studies from her home country, Brazil.

Contact: Larissa Buss, Concordia University, 1455 de Maisonneuve Blvd. W., Montreal, Quebec, Canada, H3G 1M8, (438) 823-7128, larissacks@gmail.com

REFERENCES

Baker, A. A. (2011). Discourse prosody and teachers' stated beliefs and practices. *TESOL Journal*, *2*(3), 263–292.

Baker, A., & Murphy, J. (2011). Knowledge base of pronunciation teaching: Staking out the territory. *TESL Canada Journal*, *28*(2), 29–50.

Borg, S. (2009). English language teachers' conceptions of research. *Applied Linguistics*, *30*(3), 358–388.

Breitkreutz, J., Derwing, T., & Rossiter, M. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal*, *19*(1), 51–61.

Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System*, *28*(2), 191–215.

Celce-Murcia, M., Brinton, D., Goodwin, J., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). New York: Cambridge University Press.

Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.

Derwing, T. (2008). Curriculum issues in teaching pronunciation to second language learners. In J. Edwards & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 347–369). Philadelphia, PA: John Benjamins.

Foote, J., Holtby, A., & Derwing, T. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, *29*(1), 1–22.

Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Wanier-Klimczak, E. (2012). The English pronunciation teaching in Europe survey: Selected results. *Research in Language*, *10*(1), 5–27.

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

Macdonald, S. (2002). Pronunciation: Views and practices of reluctant teachers. *Prospect*, *17*(3), 3–18.

Sifakis, N., & Sougari, A. (2005). Pronunciation issues and EIL pedagogy in the periphery: A survey of Greek state school teachers' beliefs. *TESOL Quarterly*, *39*(3), 467–488.

Tergujeff, E. (2012a). The English Pronunciation Teaching in Europe Survey: Finland. *Apples*, *6*(1), 29–45.

Tergujeff, E. (2012b). English pronunciation teaching: Four case studies from Finland. *Journal of Language Teaching and Research*, *3*(4), 599–607.

Timmis, I. (2002). Native-speaker norms and international English: a classroom view. *ELT journal*, *56*(3), 240–249.

INTEGRATED SPEAKING, LISTENING AND PRONUNCIATION: ARE TEXTBOOKS LEADING THE WAY?

Patricia Watts, University of Illinois at Urbana-Champaign

Amanda Huensch, University of South Florida

After undergoing a period of de-emphasis in the early years of communicative language teaching, pronunciation instruction is experiencing a renaissance. While access to stand-alone pronunciation courses is still somewhat limited, a growing number of learners are able to receive pronunciation instruction through integrated skills courses, such as those blending pronunciation instruction with speaking and listening. The development of this trend has been marked by a concurrent rise in integrated skills textbooks focusing on speaking, pronunciation, and listening (SPL). We examined eleven sets of SPL textbooks to determine how this integration was being achieved and how textbooks were dealing with certain inherent challenges such as finding an appropriate balance among the skills, providing guidance for inexperienced pronunciation teachers, and addressing the diverse pronunciation needs of learners from different backgrounds and levels. To that end, we documented topics covered, their relation to findings on intelligibility and comprehensibility, and support provided for novice pronunciation teachers both in terms of presenting materials and assessing performance. While results indicated several salient concerns with these books in terms of support provided to instructors and the systematic presentation of topics, they also showed that textbook syllabi primarily reflect the literature on intelligibility and comprehensibility. Recommendations for future SPL textbooks and the SPL movement are provided.

INTRODUCTION

Integrated teaching of speaking, pronunciation, and listening (SPL) has remained a topic of interest for several decades. This movement began to take shape in the 1980s with language teaching practitioners and researchers advocating for the linking of pronunciation practice with listening instruction (Gilbert, 1987) and with broader and more communicative speaking activities (Acton, 1984; Celce-Murcia, 1987; Pica, 1984). Literature outlining ideas and frameworks for integrating these skills (Levis & Grant, 2003; Murphy, 1991; Morley, 1994) helped advance the trend by offering specific suggestions for implementation. In 2003, the Speaking and Pronunciation Interest Section (SPIS) of TESOL fostered oral-aural skills integration further by adding listening to its focus, officially changing its name to SPLIS (Speaking, Pronunciation and Listening) and adopting the following as one of its six explicitly stated goals: To "encourage and support the integrated teaching of pronunciation, speaking, listening, and nonverbal communication" (SPLIS web site). Part of the appeal of this movement is based on the premise that a natural interdependence of these skills exists both in daily life and in the development of oral proficiency (Murphy, 1991), a claim that has a sound theoretical foundation and makes sense intuitively. Another factor contributing to the growing SPL

movement is the fact that many language programs do not offer stand-alone pronunciation courses, leaving this important part of language instruction to be covered alongside related skills when it is covered at all (Foote, Holtby & Derwing, 2011). In spite of growing interest, a number of concerns surround the trend, such as achieving an appropriate balance among the skills (Foote, Trofimovich, Collins & Soler Urzúa, 2013; Levis & Grant, 2003), the need for teacher expertise in teaching pronunciation (Breitkreutz, Derwing & Rossiter, 2001; Foote et al., 2011), and uncertainty about how to address diverse pronunciation needs of learners from different language backgrounds and levels of intelligibility (Foote et al., 2013; Piske, MacKay & Flege, 2001). Factors such as these complicate the task of syllabus design and materials development as well as the actual teaching of pronunciation. Despite these challenges, a number of textbooks have begun to integrate pronunciation with other oral skills. To investigate these concerns and document the specifics of how integration is being achieved, we undertook a widespread review of SPL textbooks, focusing on the following questions:

- What pronunciation topics are included and how well do syllabus selections reflect findings from the literature on intelligibility and comprehensibility?

- How does treatment of the topics stack up in terms of efficiency and thoroughness?

- What support is available for teachers who may have little or no training in teaching pronunciation?

METHOD

To select books for inclusion in the study, we first decided to limit our focus strictly to SPL books and not to include four-skills texts. This decision was driven by our specific interest in the integration of content for the three related oral-aural skills. With this focus in mind, we solicited textbook recommendations from TESOL's SPL interest section online discussion board, from colleagues with related research and teaching interests, and from representatives at major publishing houses. Based on these recommendations, eleven sets of textbooks were reviewed (see Appendix A) along with corresponding teacher's manuals and other accompanying materials and resources, such as CD ROMs and websites. Thus, when we refer to a textbook, we are including multiple levels of textbooks (if applicable) as well as teacher's manuals and any other supporting materials.

We created a template for textbook review (see Appendix B) to guide our examination based on our research questions. The template originally included the following six questions:

1. What pronunciation topics are covered?
2. What is the rationale for the selection of pronunciation content?
3. Does the pronunciation content correspond with findings from research on intelligibility?
4. How is a pronunciation focus integrated into listening and speaking assignments?
5. What guidelines or rubrics are given to assess pronunciation performance?
6. Are resources provided to aid inexperienced pronunciation teachers?

After deciding upon these six questions, the researchers randomly selected and analyzed one set of textbooks each (or approximately 10% of the data). This was done both to check the choice of template questions as well as to compare the analyses of the two researchers. The textbooks and teacher's manuals were reviewed page by page. In terms of listing the topics covered, all topics that appeared in the book were listed even if the treatment was cursory. For the questions related to the qualitative analysis, the researchers summarized their findings as well as listed example passages and activities from the textbooks as justification. After conducting separate analyses on these two sets of textbooks, the researchers compared their analyses to ensure agreement. At this stage the template of questions was also evaluated and both researchers agreed to the addition of the following three questions based on the notes from their analyses:

1. What is the overall gist/focus of the book?

2. How complete is the coverage of topics in terms of accuracy, efficiency, and thoroughness?

3. What stands out about this book (both positives and negatives)?

The remaining nine textbooks were then divided between the researchers for analysis. Once the textbook reviews were completed, the quantitative analysis was completed by tallying pronunciation topics. The qualitative analysis was conducted by combining the notes for each of the questions and comparing them for trends found across the textbooks. Each researcher completed the qualitative analysis separately after which the researchers met to compare their notes and summarize the findings.

RESULTS

What pronunciation topics are covered in the text and how well do they correspond with research findings on intelligibility and comprehensibility?

In this section, we present the results related to the pronunciation topics covered in each of the textbooks and how well they correspond with findings from the intelligibility and comprehensibility research. Topics were divided into three main sub-groups by the researchers: phrase level topics (e.g., focal stress, linking), word level topics (e.g., compound nouns, word stress) and sound level topics (e.g., /θ/ vs. /ð/, diphthongs). Table 1 below provides a summary of the topics found in each text/set of texts. An 'X' in a column indicates that the topic was presented. Note, however, that an indication that a topic was covered in no way relates to the thoroughness of coverage. A discussion of the efficiency and thoroughness of topic selections can be found in the next section.

Table 1: Pronunciation Topics Covered in the Textbooks

	All Clear: 1 & 2	Basic Tactics	Open Forum: 1 & 2	Person to Person 2	Q Skills 1, 2, 3 for Success	Real Talk	Speaking Solutions	Talk it Through 2	Tuning In	What I Believe	World English 2
Phrase											
Focal Stress		X	X	X	X	X		X		X	X
Contractions	X	X	X		X	X				X	X
Reductions	X				X	X			X	X	
Thought Groups	X				X	X		X	X	X	X
Linking	X	X	X	X	X	X	X	X		X	X
Intonation	X	X	X	X	X	X	X	X	X	X	X
Rhythm	X	X	X	X	X	X	X	X	X	X	X
h-elision		X	X			X					
Blending		X				X					
can/can't[1]			X					X	X	X	
Word											
Multi-word Verbs	X					X					
Compound Nouns	X	X			X	X			X	X	
-s endings	X	X	X		X	X	X	X	X		
-ed endings	X	X	X	X	X	X	X	X	X	X	X
Word stress		X	X	X	X	X	X	X			
Number stress		X			X						
Sound Level											
C/V Chart	X										
/iy/ vs /i/			X								
Voiced vs. voiceless Cs			X								
Theta/Eth				X		X		X			
/r/ after vowels				X							
Consonant clusters				X				X			
/ai, iy, i/									X		
Diphthongs								X	X		
/r/ vs. /l/								X			

[1] Even though *can* and *can't* are single words, they are treated as phrase-level pronunciation topics because in the stream of speech, the phonological distinction between their vowel quality is a result of differences manifested in rhythm in phrasal contexts. Unstressed *can* is subject to vowel reduction while this is not possible for *can't* because the word receives some degree of stress due to its contracted negative form.

Deciding which pronunciation topics to include is an important task for textbook authors, and the factors authors take into consideration during the decision making process warrants further exploration. The findings from intelligibility and comprehensibility research must be considered, especially because of the shift away from accent reduction in pronunciation teaching (Levis, 2005). At the phrase level, the importance of suprasegmentals in general (Derwing & Munro, 1997; Munro & Derwing, 1995; Suenobu, Kanzakei & Yamane, 1992; Zielinski, 2006) and primary stress in particular (Hahn, 2004) have been noted. These findings are reflected well in the syllabi of the books we critiqued. Notably, all of the books treat rhythmic alternations and intonation, ten books (91%) cover linking—a prevalent means of enhancing rhythm—and eight books (73%) include focal stress.

At the word level, several studies (Benrabah, 1997; Bond & Small, 1983; Zielinski, 2008) noted a correlation between word stress and intelligibility. Seven of the books reviewed (64%) included a focus on word stress, which indicates recognition of the importance of this topic. Another interesting finding is the widespread coverage of morphological endings (–ed 100%, –s, 73%). Although the contribution of morphological endings to intelligibility and comprehensibility has not been addressed, their inclusion may be warranted because they are frequent in speech, problematic among learners, and represent a potentially stigmatizing error.

Many of the studies cited previously (Derwing & Munro, 1997; Munro & Derwing, 1995; Suenobu et al., 1992; Zielinski, 2006) and Zielinski (2008) also discussed the impact of non-standard segmentals on comprehensibility and intelligibility. Overall, sound level topics received very little space on the syllabi. Six of the textbooks failed to include any coverage of segmentals, although one of these books did include a consonant and vowel chart with no explanation or corresponding activities. For the remaining five books, coverage varied from one to four lessons focusing on consonants and vowels.

Choosing which segmentals to focus on is another question textbook authors face. One means of making selections with some research basis is using functional load, which is determined by taking into account frequency of the sound or minimal pair, position of the sound in the word among other factors (King, 1967). Munro & Derwing (2006) found that non-standard segmentals with high functional load negatively impacted comprehensibility. To examine the consonants and vowels included, King's ten point scale of functional load was used as the basis for determining a load rating, with segments listed in the top five categories corresponding to high functional load and segments in the bottom five categories as low functional load. Five of the six books containing segmentals contained at least one vowel (/iy/, /ay/) or consonant (/r/, /l/) with high functional load. Yet, with the inclusion of so few segmentals, it was not really possible to make any conclusions with regard to functional load.
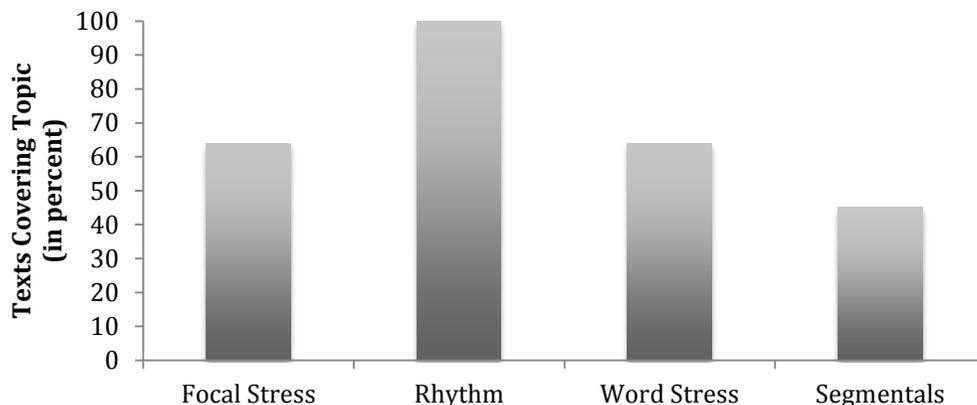
Figure 1: Syllabus Foci

Figure 1 displays information about the percent of textbooks covering topics related to intelligibility and comprehensibility. In sum, the pronunciation syllabi corresponded well with findings from the intelligibility and comprehensibility at the phrase level, somewhat at the word level, and not very well at the sound level due to the inclusion of so few segmentals.

How complete is the coverage of topics in terms of thoroughness and efficiency?

As we saw in the previous section, textbooks covered a wide variety of topics; however, the quality of coverage in terms of both thoroughness and efficiency was not considered in that discussion. During the preliminary analysis which included two sets of textbooks, it became clear that simply marking a topic as being covered or not did not provide a detailed enough picture for the differences among texts that were found in terms of coverage. Thus, a question was added to the template regarding the thoroughness and efficiency of topic coverage. When considering thoroughness, for example, we saw a tendency of a majority of texts to provide a 1-3 sentence explanation of a given topic followed by a handful of listen and repeat practice items as in the following explanation of focal stress from *Open Forum 3*:

> The focus word is the most important word in a statement. Speakers emphasize focus words by stressing them. This makes the important words easier to hear and understand. Focus words are usually content words (nouns, verbs, adverbs, and adjectives) (p. 81).

Some textbooks stood out by having a more complete explanation of the feature as well as providing practice in prediction, perception and production. The presentation of focal stress in *Talk It Through*, for example, begins with full page explanation of focal stress, connecting to a discussion of rhythm from several chapters earlier. Next, the text goes on to provide predictive rules (e.g., emphasize new information in a sentence). The explanations and predictive rules are followed by exercises in which students predict and then practice perceiving focal stress, analyzing any mistakes they made. In the next exercise, students produce marked examples with a partner and, in a final exercise, they attempt communicative practice in a group work activity. As we can see, the presentation of focal stress in *Talk It Through* goes beyond a 1-3 sentence explanation.

In terms of efficiency, it became clear that some textbooks did not take advantage of opportunities to provide general patterns of features, but instead gave only a handful of specific examples. In this way, only part of a system was presented without mention of the larger whole. For example, the focus of the pronunciation topic in Unit 7 of *Expanding Tactics for Listening* is the reduction of *is* and *are*. Rather than focusing on the use of contractions in general, or even the reduction of all *to be* verbs, this text only gave two specific examples. Similarly, in a focus on word stress, *Open Forum 1* states: "Multi-syllable words have main stress on one syllable. This stress goes on the syllable before the suffix with these suffixes: *-ogy, -ogist, ion, ity*" (p. 40). In these cases, we consider the coverage of topics to be inefficient in that opportunities for presenting general patterns of rules were not taken.

More problematic, however, was an almost universal lack of systematic coverage of topics. Most textbooks did not make explicit the connection between related topics so that learners could see the interconnected system. For example, most textbooks presented the different features of rhythm (e.g., linking, blending, reductions), in isolated sections and no attempt was made to group related topics together as part of a systematic presentation of phrase rhythm. Thus, even though in this case a thorough coverage of topics was achieved, we believe that without explicit guidance, learners would not grasp how the combination of features work in consort to create the full system. One notable exception to this finding was the textbook *Talk It Through* which, as mentioned at the beginning of this section, tended to link new topics with those previously covered in the text.

Are resources or tips provided to aid inexperienced teachers?

In this section we present results related to whether textbooks provided additional resources or tips to aid inexperienced teachers. This question was included because, as noted in Foote et al. (2011), access to teacher training in pronunciation is more often received at conference presentations and workshops rather than as a credit course in a MATESL program and teachers comment on a feeling of a lack of preparation of teaching these topics. After our initial analysis, textbooks were placed into three categories: *No*, *Limited* and *Yes*. Figure 2 below provides a summary of the categorization of texts.
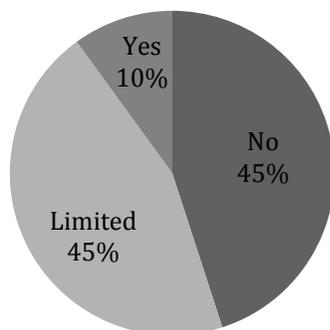

Figure 2: Tips and Resources Provided to Teachers in the Textbooks

Five of eleven or 45% of textbooks received a rating of *No* meaning they provided no other information to teachers beyond what was included in the student text. Five of eleven, or 45%, of textbooks received a rating of *Limited*. While textbooks in this category provided some additional information to teachers, it was often vague or incomplete. For example, when introducing rhythm the teacher's edition of *World English 2* provided this additional information:

> Introduce the idea of content words (meaning words) and function words (grammar words). If necessary, review the names of the part of speech and elicit more examples for each. Explain that content words have great stress (sound stronger) in a sentence (p. 78).

As we can see in this example, an attempt was made to clarify the labels 'content words', 'function words' and 'stress'; however, if an instructor did not know that 'sounding stronger' was the result of an increase in pitch, length and volume, then this limited information might not provide the intended help.

Only one textbook of eleven, *Real Talk*, received a rating of *Yes* in relation to this question. For comparison purposes, let us consider the help section related to rhythm from the teacher's manual.

> It may require a lot of ear training before students learn to hear the variations in pitch, volume, and intonation that characterize stressed words. This lesson is only an introduction. Point out to students that listening for stressed words is one of the most useful and more crucial strategies they must develop in order to understand spoken English. Refer back to this point frequently (p.17).

As we can see from this example, not only have the authors included the correlates of stress, but they have also indicated the importance of listening for stress to comprehension and reminded teachers to refer back to this point during future lessons.

Another consideration related to resources provided to teachers deals with support for the evaluation and assessment of pronunciation performance. As stated earlier, the lack of sufficient teacher training for pronunciation instruction also implies the necessity for guidance in evaluation as well. Of the eleven textbooks that we reviewed only five included any guidance on assessment in terms of either what to assess or how to assess it. Oftentimes, this information would be as simple as directing teachers to circulate around the classroom while students are working in groups and listen for accuracy on the pronunciation topic in focus. Therefore, we recommend that future textbooks consider including guidance for teachers in the area of assessment and evaluation.

DISCUSSION

Are textbooks leading the way?

While the number of SPL textbooks on the market is impressive in terms of diversity and uniqueness among the choices available, several salient concerns with the majority of these books prevent us from concluding that textbooks are leading the way. Chief among our concerns is a lack of systematic coverage of pronunciation topics, an issue that is even more important because of the possibility that some teachers may lack training in phonology and the teaching of pronunciation as has been noted in the literature (Burns, 2006; Foote et al., 2011). First and foremost, we recommend a more explicit connection

of topics throughout the text. For example, topics such as linking, trimming, blending, and reductions, should be identified as belonging to a category of features that aid in the creation of accurate phrase rhythm. This will allow students a better grasp of the entire system. Second, we feel that the potential of SPL can be realized only if the pronunciation content is truly integrated with the broader listening and speaking material rather than being presented and practiced in isolation. Such integration might include activities in which students listen and mark focal stress in a portion of a listening passage to which they have listened to previously for the thematic content or having students determine the appropriate intonation pattern of questions they will later use to interview a classmate. Three additional improvements would further strengthen syllabus design:

1) Consistently introduce basic predictive rules for all pronunciation topics so that students can use that information to guide their speaking;

2) Broaden the focus on segmentals using a principled means to select these sounds. Munro and Derwing (2006) and Levis and Cortez (2008) offer advice in that regard;

3) Create a wider range of pronunciation task types instead of relying on listen and repeat activities.

Finally, we recommend that books offer more support and tips for inexperienced teachers through teacher's manuals or websites. One way to improve this shortcoming would be to involve pronunciation experts as authors or consultants. Along with this suggestion, we concur with Derwing, Diepenbroeck, and Foote's (2012) recommendation that the language teaching profession needs to increase opportunities for teachers to develop their knowledge base and skill for teaching pronunciation.

In spite of these criticisms, we still believe that an integrated SPL curriculum can offer a meaningful way to contextualize pronunciation points in addition to providing opportunities for students to connect the various pronunciation topics covered. Two books that we reviewed bear special mention for their efforts in this regard. The first is *Real Talk*, which scored highly in all rated areas and offers authentic listening passages from a variety of sources (conversations, lectures, and interviews) to contextualize pronunciation points. We recommend the second book, *Talk it Through,* on the basis that it provides one of the most efficient and thorough treatments of pronunciation topics.

In closing, we posit that integrated SPL teaching and materials development is at a critical juncture. Along with the proposed changes to textbooks, we also recommend investigation of SPL classroom practices in the form of pedagogically oriented research to gain additional information about how best to integrate all oral-aural skills in a single course. These steps will help infuse the movement with vitality and lead to greater effectiveness.

ACKNOWLEDGEMENTS

ABOUT THE AUTHORS

Patricia Watts is the Coordinator of the International Teaching Assistant program at the University of Illinois at Urbana-Champaign. Her areas of interest include ITA training, oral fluency, materials development, and technological applications for the teaching of pronunciation.

Amanda Huensch earned her Ph.D. and MATESL from the University of Illinois at Urbana-Champaign. She is currently an Assistant Professor of Applied Linguistics in the Department of World Languages at the University of South Florida. Her research interests include the acquisition of second language phonology and improving ESL/TESL pedagogy.

Patricia Watts, University of Illinois at Urbana-Champaign, 4080 Foreign Languages Building, 707 S. Mathews Ave., Urbana, IL 61801, (217) 333-1506, pawatts1@illinois.edu

Amanda Huensch, University of South Florida, CPR 419, 4202 E. Fowler Ave., Tampa, FL, 33620, (813) 974-2548, huensch@usf.edu

REFERENCES

Acton, W. (1984) Changing fossilized pronunciation. *TESOL Quarterly, 18,* 71-86.

Benrabah, M. (1997). Word-stress: A source of unintelligibility in English. *International Review of Applied Linguistics*, *35*, 157–165.

Bond, Z., & Small, L. (1983). Voicing, vowel, and stress mispronunciations in continuous speech. *Attention, Perception, and Psychophysics, 34*, 470-474.

Breitkreutz, J., Derwing, T. M., & Rossiter, M. J. (2002). Pronunciation teaching practices in Canada. *TESL Canada Journal, 19*, 51-61.

Burns, A. (2006). Integrating research and professional development on pronunciation teaching in a national adult ESL program. *TESL Reporter, 39,* 34-41.

Celce-Murcia, M. (1987). Teaching pronunciation as communication. In J. Morley (Ed.), *Current Perspectives on Pronunciation* (pp. 29-40). Washington, D.C., TESOL.

Derwing, T., Diepenbroek, L. G., & Foote, J. A. (2012). How well do general skills ESL textbooks address pronunciation? *TESL Canada Journal, 30,* 22-44.

Derwing, T., & Munro, M. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition, 19,* 1-16.

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of pronunciation teaching in adult ESL programs in Canada, 2010. *TESL Canada Journal*, *29*, 1- 22.

Foote, J. A., Trofimovich, P., Collins, L., & Soler Urzúa, F. (2013). Pronunciation teaching practices in communicative ESL classes. *The Language Learning Journal*. Published online 16 April 2013. doi:10.1080/09571736.2013.784345

Gilbert, J. (1987). Pronunciation and listening comprehension. In J. Morley (Ed.), *Current Perspectives on Pronunciation* (pp. 29-40). Alexandria, VA, TESOL.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38,* 201-223.

King, R. D. (1967). A measure for functional load. *Studia Linguistica, 21,* 1-14.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39,* 369-377.

Levis, J., & Grant, L. (2003). Integrating pronunciation into ESL/EFL classrooms. *TESOL Journal, 12,* 13-19.

Levis, J., & Cortes, V. (2008). Minimal pairs in spoken corpora: Implications for pronunciation assessment and teaching. In C. A. Chapelle, Y.-R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 197-208). Ames, IA: Iowa State University.

Morley, J. (1994). A multidimensional curriculum design for speech-pronunciation instruction. In J. Morley (Ed.). *Pronunciation Pedagogy and Theory: New Views, New Directions* (pp. 64-91). Alexandria, VA: TESOL.

Munro, M., & Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45,* 73-97.

Munro, M., & Derwing, T. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34,* 520-531.

Murphy, J. (1991). Oral communication in TESOL: Integrating speaking, listening, and pronunciation. *TESOL Quarterly, 25,* 51-75.

Pica, T. (1984). Pronunciation activities with an accent on communication. *English Teaching Forum, 22,* 2-6.

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics 29*, 191-215.

SPLIS website (n.d.). Retrieved July 11, 2013, from http://www.tesol.org/connect/interest-sections/speech-pronunciation-and-listening

Suenobu, M., Kanzakei, K., & Yamane, S. (1992). An experimental study of intelligibility of Japanese English. *International Review of Applied Linguistics, 30,* 146-153.

Zielinski, B. W. (2006). The intelligibility cocktail: An interaction between speaker and listener ingredients. *Prospect, 21,* 2-45.

Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System, 36,* 68-84.

Appendix A: Textbooks and Teacher's Manuals Surveyed

Baker, L. & Tanka, J. (2006). *Real talk 1: Authentic English in context.* New York: Pearson-Longman.

Baker, L. & Tanka, J. (2006). *Real talk 2: Authentic English in context.* New York: Pearson-Longman.

Baker, L. & Tanka, J. (2006). *Real talk 1: Authentic English in context, teacher's manual.* New York: Pearson-Longman.

Blackwell, A. & Naber, T. (2006). *Open forum 1: Academic listening and speaking.* New York: Oxford University Press.

Blackwell, A. & Naber, T. (2006). *Open forum 2: Academic listening and speaking.* New York: Oxford University Press.

Böttcher, E. (2008). *What I believe: Listening and speaking about what really matters.* NewYork: Pearson-Longman.

Brooks, M. (2011). *Q 2: Skills for success listening and speaking*. New York: Oxford University Press.

Craven, M. & Sherman, K. D. (2011). *Q 3: Skills for success listening and speaking*. New York: Oxford University Press.

Duncan, J. & Parker, A. (2007). *Open forum 3: Academic listening and speaking.* New York: Oxford University Press.

Fragiadakis, H. K. (2007). *All clear 2: Listening and speaking, 3rd edition*. Boston: Heinle.

Fragiadakis, H. K. (2008). *All clear 1: Listening and speaking, 2nd edition*. Boston: Heinle.

Johannsen, K. L. & Chase, R. T. (2010). *World English 2: Real people, real places, real language.* Boston: Heinle.

Johannsen, K. L. & Chase, R. T. (2010). *World English 2: Real people, real places, real language: teacher's edition.* Boston: Heinle.

Kozyrev, J. R. & Baker, M. (2001). *Talk it through 2: Listening, speaking, and pronunciation*. New York: Houghton Mifflin.

Matthews, C. (1994). *Speaking solutions: Interaction, presentation, listening, and pronunciation skills.* New York: Pearson-Longman.

Matthews, C. (1994). *Speaking solutions: Interaction, presentation, listening, and pronunciation skills, instructor's manual.* New York: Pearson-Longman.

Numrich, C. (2006). *Tuning in: Listening and speaking in the real world*. New York: Pearson-Longman.

Richards, J., Bycina, D., & Wisniewska, I. (2005). *Person to person 2: Communicative speaking and listening skills, 3rd edition.* New York: Oxford University Press.

Richards, J., Bycina, D., & Wisniewska, I., & Kocienda ,G. (2006). *Person to person 2: Communicative speaking and listening skills, 3rd edition.* New York: Oxford University Press.

Richards, J. & Trew, G. (2011). *Basic tactics for listening, 3rd edition*. New York: Oxford University Press.

Richards, J. & Trew, G. (2011). *Basic tactics for listening, teacher's book, 3rd edition*. New York: Oxford University Press.

Richards, J. & Trew, G. (2011). *Developing tactics for listening, 3rd edition*. New York: Oxford University Press.

Richards, J. & Trew, G. (2011). *Developing tactics for listening, teacher's book, 3rd edition*. New York: Oxford University Press.

Richards, J. & Trew, G. (2001). *Expanding tactics for listening, 3rd edition.* New York: Oxford University Press.

Richards, J. & Trew, G. (2001). *Expanding tactics for listening, teacher's book, 3rd edition.* New York: Oxford University Press.

Scanlon, J. (2011). *Q 1: Skills for success listening and speaking*. New York: Oxford University Press.

Appendix B: Template for Textbook Review

Title:

Publisher:

Date of Publication:

Level:

1. What is the overall gist/focus of the book?

2. What is the rationale for the selection of pronunciation content?

3. What pronunciation topics are covered?

4. How is the coverage of topics in terms of accuracy, efficiency, and thoroughness?

5. Does the pronunciation content correspond with findings from research on intelligibility?

6. How is a pronunciation focus integrated into listening and speaking assignments?

7. Are guidelines and rubrics given to evaluate/assess pronunciation performance?

8. Are resources or tips provided to aid inexperienced pronunciation teachers?

9. What stands out about this book (both positives and negatives)?

de Castro Gomes, M. L. (2013). Understanding the Brazilian way of speaking English. In Levis, J. & LeVelle, K. (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference,* Aug. 2012. (pp. 279-289). Ames, IA: Iowa State University.

UNDERSTANDING THE BRAZILIAN WAY OF SPEAKING ENGLISH

Maria Lúcia de Castro Gomes, Universidade Tecnológica Federal do Paraná

To better understand the aspects that characterize Brazilians' pronunciation of English, we have formed a study group, formed by researchers, professors, and under-graduate students, to study phonetics, phonology and pronunciation teaching, with a theoretical focus on usage-based phonology (Bybee, 2001, 2010) and exemplar models (Pierrehumbert, 2000, 2001, 2003). The group has had an interesting experience using Praat (Boersma & Weenink), a software for acoustic analysis of speech data. The objective of this article is to demonstrate the basis for the formation of the group and to show how the participants have been using their own speech data, recorded in our laboratory, to understand the acquisition of English phonology by Brazilians. As a result of the experience, group members have improved their own pronunciation, developed educational activities to use in teaching pronunciation, and carried out their own research projects.

INTRODUCTION

Foreign language teacher development should require a solid training in phonetics and phonology for the effective teaching of pronunciation. English teachers, especially in light of the role of English as an international language, should be prepared to teach their students not only to produce the sounds of the language in an intelligible form, but also to understand the speech of speakers from many backgrounds. There was a time when the choice for the teaching / learning of English was limited to the varieties of American or British English. Today the focus is on English as a Lingua Franca and the possibility of communication with English users around the world.

In this context, the university must be the locus for discussion about this new paradigm, both between teachers and students, and with teachers from public and private schools. The discussion should not be limited to theoretical and philosophical questions about language acquisition and teaching, but should prioritize discussion of pronunciation teaching practices within this new perspective.

Also important for teacher development is to be a member of an association dedicated to the development of the professional activity. English teachers in Brazil have BRAZ-TESOL, the largest association of teachers of English to speakers of other languages in the country. Founded in 1986, BRAZ-TESOL is a nonprofit organization with over 2,000 members and is affiliated with TESOL International USA, the UK IATEFL, and is a member of the Southern Cone TESOL. In Curitiba, our city, teachers can count on the BRAZ-TESOL Regional Chapter, which in addition to promoting events for the development of English teachers, also organizes special interest groups in several areas related to the teaching of English.

Whereas the main concern of an association like BRAZ-TESOL is with the improvement of teaching practices, the main concern of the university is with the theoretical basis for the professional activity. Having that in mind, we proposed a partnership between UTFPR - Federal Technological University of Parana, specifically through DALEM, Department of Modern Foreign Languages, and BRAZ-TESOL, through its Regional Chapter Curitiba, more specifically its Pronunciation RIG – Regional Interest Group, for the formation of a study group on phonetics and phonology for the teaching of English.

Composed of university professors, English teachers from public and private schools, and students, the group laid the groundwork for its studies in the first meeting. Firstly, teachers and students are registered in a CNPq[1] Research Group. The theoretical bases for studies from the group were to come from three fronts: a probabilistic view for language acquisition, English as a Lingua Franca, and acoustic phonetics for instrumental analysis. From these early definitions, the group established and defined its operation. This article presents the theoretical concepts that underlie these studies, the use of PRAAT as a tool for understanding the main characteristics of Brazilians speaking English, and the research topics selected by the group.

The Theoretical Bases

Three models were chosen as a framework for the group studies: Probabilistic Linguistics, Usage-based Phonology and Exemplar Model. Frequency of use, gradience of categories and multirepresentation of forms are coincident aspects of language that permeate the three models.

According to Bod, Hay & Janedy (2003), language shows evidence of a probabilistic system. All levels of representation in phonetics and phonology show statistical variation and speakers have implicit knowledge of language change (Pierrehumbert, 2003). In usage-based phonology, grammar is seen as the cognitive organization of the experience a speaker has with language (Bybee, 2006). In this model, three points are crucial: the creative role of repetition, the effects of frequency and the emerging character of grammar. Following the same line of thought, in the Exemplar Model there are also three fundamental bases: phonetic details, gradience of mental representations, and the frequency. As a result of contact with the language, the speaker will form an exemplar map, forming clouds, and taking into account social, pragmatic, semantic, morphological, phonological and phonetic factors (Pierrehumbert, 2000, 2001).

As can be seen, the three models are formed from the same pillars: the effects of frequency, the gradience of linguistic forms, and the dynamic character of language. This in itself would justify the choice of the theoretical basis. However, another choice for the studies, the perspective of English as a Lingua Franca, also was seen as a good match for the assumptions of those three models.

The Perspective of Teaching / Learning English as a Lingua Franca

Given the expansion of English, it is said that the language does not belong to the people that have English as their native language anymore, because the number of non-native speakers is double the combined population of the English speaking countries (Crystal, 2010). Furthermore, communication in English without the presence of a native speaker is an increasing probability. According to Jenkins (2000), speakers of English as an L1 no longer have the right to dictate the

---

[1] CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico – National Counsel of Technological and Scientific Development.

standards of pronunciation for its use as an L2. Having that in mind, what would the ideal pattern be?

Jenkins proposes an ideal model for teaching pronunciation of English as an international language: that of fluent bilingual (non-native) speakers, as they would be more realistic and appropriate models than speakers of English as an L1 (Jenkins, 2000). From her research, the author established a series of priority items in relation to teaching English pronunciation, which she calls the Lingua Franca Core (LFC). Containing characteristics of American English, British English and varieties of English as L2, the LFC, according to Jenkins, allows for a focus on intelligibility, and she argues that specific characteristics of native varieties that are difficult for the learner to acquire should not be considered for teaching if not relevant in international communication.

Along with the theoretical basis for the studies, and the perspective of English as a lingua franca, the group intended to perform acoustic analyses of speech data.

The First Semester – Methodology

From the choices made, the group established a schedule and started the activities with a review of the sound systems of English and Portuguese. Next there was a discussion about the characteristics of English produced by Brazilian speakers. There has been a great deal of research work about Portuguese-English interphonology conducted at several universities in Brazil and abroad. Afterwards data was collected and a corpus composed of speech samples recorded by foreign visitors at the University and the group members themselves, was formed for analysis and awareness. The idea was to compare the data produced by native speakers with those produced by Brazilians to examine the phenomena described in those studies about Portuguese-English interphonology. Due to the perspective adopted by the group, it is important to mention that those phenomena were treated as special features of Brazilian speech, not as errors. Moreover in the line of English as a Lingua Franca, while priority will be given to the items of LFC[2] (Jenkins 2000), other items that are not considered by Jenkins but are part of Brazilians' characteristic speech will also be considered in the analyses.

After the composition of the corpus, there was a search for papers in which the items of the LFC are the focus. The next step was to edit sound files and compare the data produced by native English speakers and Brazilian English speakers, with the use of PRAAT[3].

Finally the members of the group, mostly in pairs or trios, chose research topics to initiate their own studies based on Probabilistic Linguistics, Usage-based Phonology and Exemplar Model, having the underlying premise that lexical frequency and experience with language are essential to the development of the cognitive structures in language acquisition.

The Lingua Franca Core and the Brazilian Way of Speaking English

According to Jenkins (2000), if a person wants to be intelligible in ELF communication, four areas are essential: individual consonant sounds, groups of consonants (clusters), vowels and nuclear stress placement. Regarding the consonants, Walker (2010, p. 29) alerts that "because of the impact the substitution of consonants has on ELF communication, the LFC requires speakers

---

[2] LFC – Lingua Franca Core – "a list of pronunciation items central to maintaining mutual intelligibility of ELF" (Walker, 2010).

[3] PRAAT is a computer program for acoustic analysis, developed by Paul Boersma and David Weenink, from the University of Amsterdam, downloadable for free at http://www.fon.hum.uva.nl/praat/, access on April, 2013.

to be competent, both receptively and productively, in all but two of the consonant phonemes of English." The exceptions are the consonants /ɵ/ and /ð̌/, sounds for the 'th'', as in *think* and *then*, and the production of dark [ɫ]. Jenkins argues that possible substitutions of the fricatives /ɵ/ and /ð̌/, and the replacement of the dark [ɫ] for the clear /l/ or /ʊ/ are unproblematic for EIL[4] intelligibility. The LFC also includes phonetic features for the optimal pronunciation of consonants, such as the voiceless plosives /p/, /t/ and /k/, which are pronounced with aspiration [ʰ] in initial position in stressed syllables. That is an aspect of English pronunciation that is important in our context because Brazilians tend not to produce the aspiration (Arantes, 2007; Cohen, 2004; Zimmer & Alves, 2007). In Figure 1, we can clearly see a larger aspiration in the first consonant of the word *car* produced by an American speaker, compared with the production of a Brazilian speaker. The highlighted portion shows the extent of the voicelessness from the aspiration (the Voice Onset Time) or lack of aspiration.

CAR
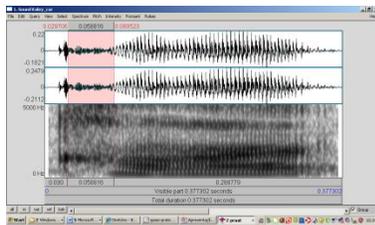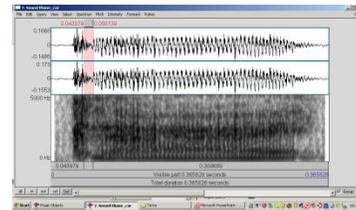
American Speaker   **Brazilian speaker**

Figure 1. Comparison of aspiration of /k/ in the word *car* by and American speaker and a Brazilian speaker.

Another special characteristics of the Brazilian way of speaking English is the palatalization of the alveolar plosives /t/ and /d/. According to Cristofaro-Silva (2011, p. 168), palatalization is the phenomenon in which a consonant acquires palatal or close to palatal articulation. In Brazilian Portuguese, palatalization of alveolar stops occurs before front high vowels or a palatal glide. The waveform and the spectrogram in Figure 2 show the palatalization in the end of the word *beside*, produced because of /i/ paragoge.

BESIDE
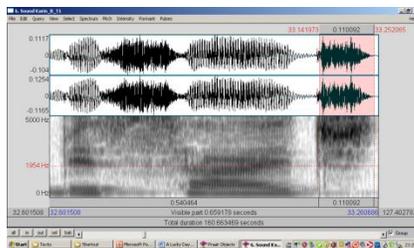
 Figure 2. Production of the word *beside* with /i/ paragoge and palatalization of the /d/, producing a final syllable [dʒi].

---

[4] EIL – English as an International Language.

Regarding consonant clusters, several researchers have analyzed the tendency of Brazilian speakers to add an epenthetic vowel to break difficult consonant clusters, especially those produced with the addition of the *–ed* morpheme to form the past tense of regular verbs (Alves, 2004; Delatorre, 2006; Gomes, 2009). Likewise there is a tendency for a paragogic vowel in words ending in consonants that are not allowed in the end of syllables in Portuguese (Cruz, 2008; Ferreira, 2007; Koerich, 2002; Silva Filho, 1998). The waveforms and spectrograms below (Figure 3) show clearly the resyllabification process performed by the Brazilians. Brazilian speaker 1 (B1 level) resyllabifies the word *passed* with an epenthetic vowel, adding an extra syllable to the word while Brazilian speaker 2 (A1 level) produces a three-syllable word with epenthesis and paragoge.

PASSED

American Speaker   **Brazilian speaker 1**   **Brazilian speaker 2**

Figure 3. Production of the word *passed* by an American speaker and two Brazilian speakers in two different levels of proficiency (B1 and A1).

As regards vowel sounds, Jenkins mentions two considerations: quality and quantity. "While vowel quantity is reasonably stable across varieties of English, vowel quality is not" (Jenkins, 2000, p. 144). This is the factor which led Jenkins to argue for a focus of attention on the long-short differences between vowels rather than on the quality of each one[5].

Rauber (2006), in her analysis of production and perception of /ɪ−i/, /ʊ-u/ and /ɛ- æ/ by Brazilians, concludes that "native-like performance seems to take place only in isolated cases, and maybe it is not possible or very unlikely to be acquired if one does not live in a country where the target language is the predominantly spoken language" (p. 149). In the waveforms and spectrograms below (Figure 4), the Brazilian speaker's vowel /i/ is longer than that of the American's. Being in a very advanced level of the language, the speaker preserves a phonetic

contrast between /ɪ/ and /i/ with exaggerated length, corroborating the dissimilation hypothesis posited by Flege, Schirru & MacKay (2003, p.471): "category dissimilation will occur only if a new L2 category is relatively close in phonetic space to a pre-existing L1 category". Comparing productions of English vowels by Italian speakers, the authors hypothesized that the early bilinguals, who seldom used Italian, would have a production with more movement than the native speakers who were monolinguals would. That prediction was supported by tokens with exaggerated movement produced by participants who were bilinguals from an early age.

---

[5] The LFC includes need for a good approximation to the native speaker quality of the central vowel / ∈ ⅂/ (Walker, 2010, p. 34).

BEAT
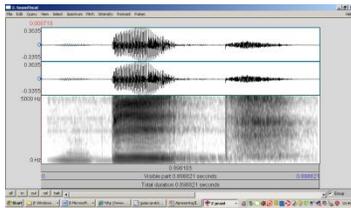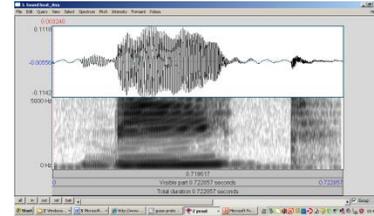
American Speaker   **Brazilian speaker**

BIT
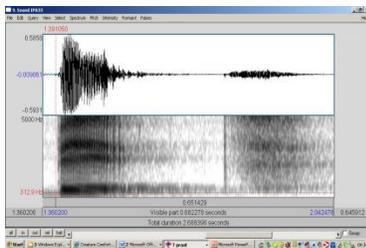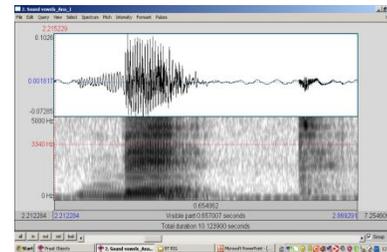
**American Speaker**   **Brazilian speaker**

Figure 4. Production of the words *beat* and *bit* by an American speaker and a Brazilian speaker in an advanced level of proficiency.

Besides differences in length at the phonological level, as in /ɪ-i/, /ʊ-u/, there are also differences in length conditioned by voicing of the following consonant. A longer vowel in English might be produced when followed by a voiced consonant than when the next consonant is a voiceless one (Kent & Read, 2002). Comparing the waveforms and the spectrograms below (Figure 5), there is more evident difference between the vowel duration of *eyes* and *ice* in the American's production than in the Brazilian's production.

EYES   **ICE**
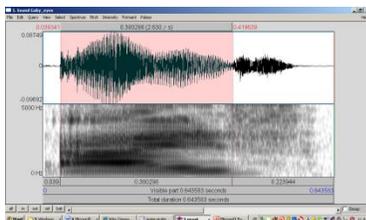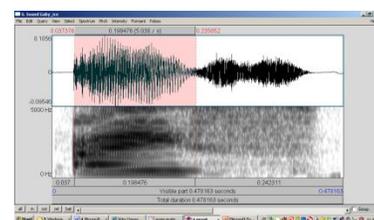
American Speaker
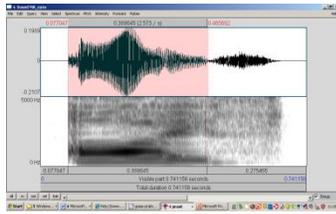
EYES

ICE
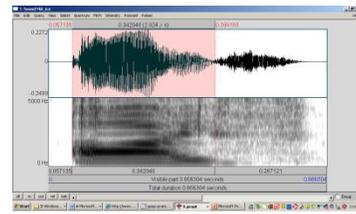
Brazilian speaker

Figure 5. Production of the words *eyes* and *ice* by an American speaker and a Brazilian speaker.

Another context for vowel duration which seems to be a difficulty for Brazilian speakers is that of final /i/ in CVCV words. Ferreira (2007) compared English native speakers' with Brazilians' production of CVCV words ending in "y", as in *daddy*. She concluded that Brazilians tend to produce a shorter final vowel when compared to native speakers. The waveforms and spectrograms in Figure 6 corroborate those results.

DADDY
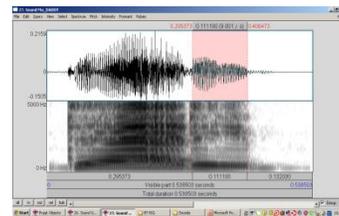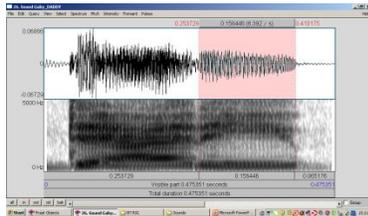
American Speaker   **Brazilian speaker**

Figure 6. Production of the word *daddy* performed by an American speaker and a Brazilian speaker.

The last of the major pronunciation issues that Jenkins lists in her LFC is nuclear stress placement and division of speech stream into word groups, because, according to the author, the correct placement of nuclear stress is important for intelligibility (Jenkins, 2000). Comparing the production of three native speakers with those of three advanced-leveled Brazilian speakers (Figure 7), there is consistent stress on the word *money,* in the middle of the sentence by the native speakers, and on the word *from* in the end of the sentence by the Brazilian speakers.

WHERE DID ALL THIS MONEY COME FROM?
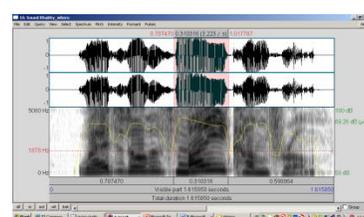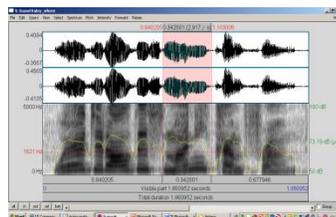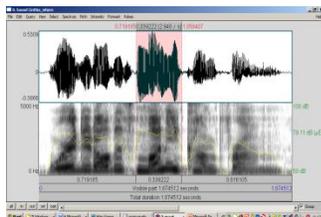
American Speaker 1   **American Speaker 2**   **British Speaker**

| Brazilian speaker 1 | Brazilian speaker 2 | Brazilian speaker 3 |

Figure 7. Comparison of the nuclear stress placement in the question *Where did all this money come from?* by three native speakers and three Brazilian speakers.

Non-Core Features and Brazilians' Difficulties

Jenkins advocates that the LFC "drastically simplifies the pedagogic task by removing from the syllabus many time-consuming items which are either unteachable or irrelevant for EIL"

(Jenkins, 2000, p. 160). These items are: /ɵ/ - /ð̆/ and final /ɫ/; exact vowel quality, pitch movement, word stress, stress-timing rhythm, vowel reduction, weak forms and the schwa vowel

/ə/, and some aspects of connected speech, such as assimilation and coalescence. Two of those are relevant when we consider Brazilian speakers of English: word stress and the production of schwa.

A good deal of work have been carried out to explain the difficulties of Brazilians with word stress in English following the studies by Baptista in the 1980s (Bertochi, 2009; Brawerman, 2006, 2012; Watkins, Brawerman & Bertochi, 2010). The waveforms and spectrograms in Figure 8 show clearly the syllable change in stress position.

## POLICE
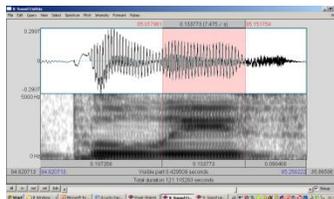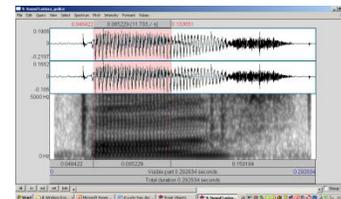
| American Speaker | Brazilian speaker |

Figure 8. Comparison of word stress placement in the word *police* by an American speaker and a Brazilian speaker.

The other aspect of Brazilian accent in English to be considered here, which is not listed in the LFC, is that of vowel reduction. According to Watkins (2001, p.5), "Brazilian speakers of English […] continue to use full vowels in some cases where a native speaker would always use a reduced one." In Figure 9, the vowel /ə/ is produced only by the American speaker. The Brazilian speaker pronounces /əʊs/ instead, probably influenced by spelling.

FAMOUS
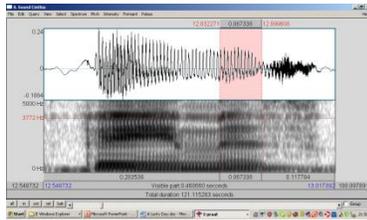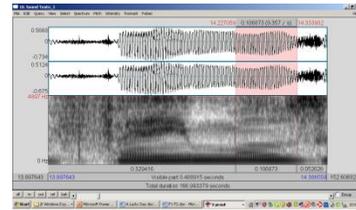
American Speaker   **Brazilian speaker**

Figure 9. Comparison of the pronunciation of the unstressed vowel in the word *famous* by a native speaker and a Brazilian speaker.

From the comparisons above the group members were able to raise awareness about the main aspects that comprise the Brazilian way of speaking English, and were also able to select those aspects that would be interesting to carry on new research projects. The focus of the projects was not only the production by Brazilian speakers, but also on the perception of their speech by other Brazilians, by other non-native speakers, and by native speakers of English, with the purpose of testing the LFC.

Research Projects

From the studies about Portuguese-English interphonology, some items were selected for

additional research. In vowel sounds the pairs /ɪ-i/, /ʊ-u/ and /ɛ- æ/ were to be investigated following the results of Rauber (2006). Another issue related to vowel sounds will be vowel length following voiced and voiceless consonants, specifically /s/ and /z/. A third study was about aspiration of the voiceless plosives /p/, /t/ and /k/, focused on perception. The next choice was the realization of words with the –*ed* morpheme. That is a very relevant feature of Brazilian English, with several research studies on the production of epenthesis. The objective was to investigate the perception by speakers of different nationalities of data produced by Brazilians in different levels of proficiency. Last but not least there was a comparison on the production of an interrogative sentence by Brazilians in different levels of proficiency and native speakers of English, with focus on nuclear stress.

FINAL REMARKS

The aim of this paper was to inform about a study group in phonetics and phonology, and the teaching of pronunciation of English as a Lingua Franca. Despite the short period of its existence, the results of their research have been presented in several events. But even more important are the future prospects with all the research in progress and the great possibility of personal and professional development of all participants.

ABOUT THE AUTHOR

Maria Lúcia de Castro Gomes is a Professor of Phonetics and Phonology of English and Portuguese at Universidade Tecnológica Federal do Paraná in Brazil. She received her doctorate in Linguistics from Universidade Federal do Paraná and the focus of research was the phenomenon of epenthesis in the pronunciation of words with the –*ed* morpheme by Brazilian speakers. Interested in teacher development, she has been voluntarily involved with BRAZ-

TESOL for over 10 years, more recently coordinating an interest group in English pronunciation teaching. Her current area of interest is forensic phonetics with some research projects under development about disguise, imitation and impersonation. Email address: malucastrogomes@terra.com.br

REFERENCES

Alves, U. K (2004). *O papel da Instrução Explícita na Aquisição Fonológica de Inglês como L2*: Evidências Fornecidas pela Teoria da Otimidade. (Unpublished MA dissertation) Universidade Católica de Pelotas, UCPel,. Pelotas, Brazil.

Arantes, V. T. P. (2007). *Perception and production of English final stops by young Brazilian EFL students*. (Unpublished MA dissertation) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Bod, R., Hay, J., & Janedy, S. (2003) *Probabilistic linguistics.* Cambridge, MA: MIT Press.

Boersma, P. & Weenink, D. Praat: doing phonetics by computer (Version 5.0.43) [Software]. Available from http://www.fon.hum.uva.nl/praat/.

Bertochi, M. M. (2009). *Padrão acentual dos compostos e sintagmas do inglês: A percepção do aprendiz brasileiro*. (Unpublished MA dissertation) Universidade Federal do Paraná, Curitiba Brazil.

Brawerman, A. (2006). *Uma análise de erros de estudantes brasileiros de inglês na acentuação de palavras com sufixos*. (Unpublished MA Dissertation) Universidade Federal do Paraná, Curitiba, Brazil.

Brawerman, A. (2012). *Os efeitos de um treinamento de percepção na aquisição do padrão acentual pré-proparoxítono da língua inglesa por estudantes brasileiros*. (Unpublished doctoral dissertation) Universidade Federal do Paraná, Curitiba, Brazil.

Bybee, J. (2001). *Phonology and language in use*. Cambridge, UK: Cambridge University Press.

Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language.* 82 (4), 711-733.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge, UK: Cambridge University Press.

Cohen, G. V. (2004). *The VOT dimension: A bi-directional experiment with English and Brazilian Portuguese stops*. (Unpublished MA dissertation) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Cristófaro-Silva, T. (2003). *Fonética e fonologia do português*. São Paulo, Brazil: Contexto, 2003.

Crua, N. C. (2008). Vowel insertion in the speech of Brazilian learners of English: a source of unintelligibility? In: Silveira, R et al (Org.). *Ilha do Desterro a Journal of English Language, Literatures in English and Cultural Studies*, 55, 133-152.

Crystal, D. (2010) *English as a global language*. 2.ed. Cambridge, UK: Cambridge University Press.

Delatorre, F. (2006). *Brazilians EFL learners'production of vowel epenthesis in words ending in –ed*. (Unpublished MA dissertation) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Ferreira, A. P. P. (2007). *Pet ou petty? Diferenças entre sequências CVC e CVCV do inglês por aprendizes brasileiros*: Uma análise acústica. (Unpublished MA dissertation) Universidade Federal do Paraná, Curitiba, Brazil.

Flege, J. E., Schirru, C. & MacKay, I. R. A. (2003) Interaction between the native and second language phonetic subsystems. *Speech communication*. 40, 467-491.

Gomes, M. L. C. (2009). *A produção de palavras do inglês com o morfema –ed por falantes brasileiros*: uma visão dinâmica. (Unpublished doctoral dissertation) Universidade Federal do Paraná, Curitiba, Brazil.

Jenkins, J. (2000). *The Phonology of English as an international language.* Oxford, UK: Oxford University Press.

Kent, R.D., & Read, C. (2002) *Acoustic analysis of speech*. 2 ed. Albany, NY: Singular – Thomson Learning.

Koerich, R.D. (2002). *Perception and production of word-final vowel epenthesis by Brazilian EFL students*. (Unpublished doctoral dissertation) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Pierrehumbert, J. B. (2000). What people know about sounds of languages. *Studies in the Linguistic Sciences*. 29, 111-120.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, J.& Hopper, P. (Eds) *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins, 137-157.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and Acquisition of phonology. *Language and Speech*. 46, (2-3), 115-154.

Rauber, A. (2006). *Perception and Production of English vowels by Brazilian EFL speakers*. (Unpublished doctoral dissertation) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Silva-Filho, J. (1998). *The production of English syllable-final consonants by Brazilian learners* (Unpublished MA dissertation) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Walker, R. (2010). *Teaching pronunciation of English as a Lingua Franca*. Oxford, UK: Oxford University Press.

Watkins, M. A. (2001). *Variability in vowel reduction by Brazilian Speakers of English*. (Unpublished doctoral dissertation) Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Watkins, M., Albini, A. B. & Bertochi, M. (2010) Suffering from stress: two English stress patterns that give Brazilian a hard time. In Rauber, A. S., Watkins, M. A., Silveira, R. & Koerich, R. D. *The Acquisition of Second Language Speech: Studies in Honor of Professor Barbara O. Baptista*. Florianópolis, Brazil: Editora Insular, 305-318.

Zimmer, M.C.& Alves, U.K. (2007) A dessonorização terminal na aprendizagem da L2: evidências do continuum fonética-fonologia. *Letras de Hoje*. 42 (3), 56-68.