PRONUNCIATION AND INTELLIGIBILITY: ISSUES IN RESEARCH AND PRACTICE

PROCEEDINGS OF THE 2ND ANNUAL PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING CONFERENCE

Editors
John Levis
Kimberly LeVelle

ACKNOWLEDGEMENTS

The papers in this volume were presented at the Pronunciation in Second Language Learning and Teaching conference held at Iowa State University on September 10 – 11 of 2010.

We thank Murray J. Munro, the invited speaker, for his enthusiastic participation. We also thank the Intensive English and Orientation Program as well as the Department of English at Iowa State University for their financial support.

2011
TESL/Applied Linguistics
Iowa State University

PRONUNCIATION AND INTELLIGIBILITY: ISSUES IN RESEARCH AND PRACTICE

Proceedings of the
2nd Pronunciation in Second Language Learning and Teaching Conference

September 10-11, 2010
Iowa State University

Acquisition and Description

Levis, J. & LeVelle, K. (2011). Pronunciation and intelligibility: An overview of the conference. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 1-6), Ames, IA: Iowa State University.

PRONUNCIATION AND INTELLIGIBILITY: AN OVERVIEW OF THE CONFERENCE

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

A generally accepted goal of pronunciation pedagogy is to help learners achieve a comfortably intelligible pronunciation rather than a native-like one. But what should this goal look like in the kinds of research studies we conduct, in our classroom practice, and in the creation of self-study and computer assisted pronunciation materials? Which elements of pronunciation are most important in achieving a comfortably intelligible pronunciation? What principles can help teachers make decisions regarding intelligibility?

The 2nd Pronunciation in Second Language Learning and Teaching Conference was held at Iowa State University on September 10-11, 2010 with the goal of exploring some of these questions. Research was presented that concerned not only English pronunciation, but also Spanish, French and Chinese pronunciation, providing a valuable forum for bringing together researchers in different areas who have common concerns. To give a flavor of the overall variety of the conference sessions, the conference schedule is reproduced below. Those sessions that are in bold are included in the Proceedings.

2010 Conference Schedule

Friday, September 10

9:00 a.m.	***Intelligibility tales*** **Laura Hahn and Patricia Watts (University of Illinois)**
9:30 a.m.	***The interlanguage speech intelligibility benefit: The case of Arabic-accented English*** **Ghazi Algethami (University of York, UK)**
10:00 a.m.	***Integrated pronunciation: Listening comprehension and intelligibility in theory and practice*** **Marnie Reed (Boston University)**
10:30 a.m.	Break
10:50 a.m.	***The pronunciation of /z/ in coda clusters in Somali-accented English*** **Ettien Koffi (St. Cloud State University)**
11:20 a.m.	*The role of word familiarity and word context on English vowel intelligibility* Ron Thomson (Brock University, Canada)
11:50 a.m.	*Don't take that tone: Critical errors in lexical tone acquisition* Alina Twist, Alison Blodgett, Jessica Bauman, Melissa Fox, Anton Ritting, Matt Winn (University of Maryland CASL)

12:20 p.m.	***Pronunciation learning strategies that improve ESL learners' linking*** **Veronica Sardegna (University of Texas at Austin)**
12:50 p.m.	Lunch
2:15 p.m.	**[Plenary]** ***Intelligibility: Buzzword or buzzworthy?*** **Murray Munro (Simon Fraser University, Canada)**
3:30 p.m.	***Destressing: Another step toward natural speech*** **Wayne Dickerson (University of Illinois)**
4:00 p.m.	***Pronunciation pedagogy: An exploration of teacher learning and classroom practices*** **Amanda Baker (Georgia State University)**
4:30 p.m.	*Towards more effective L2 pronunciation instruction: Maximizing accuracy and oral fluency* Bertha Chela-Flores (Universidad Simon Bolivar, Venezuela)
6:30-9:00 p.m.	Conference dinner

Saturday, September 11

9:00 a.m.	***The role of technology in pronunciation teaching: The use of biovisual feedback in improving Vietnamese learners' production of English syllable margins*** **Lucy Pickering (Texas A & M), Pamela Pearson & Rachel Da Silva (Georgia State University**)
9:30 a.m.	*Where is intelligibility in machine recognition of non-native speech?* Brian Teaman (Osaka JoGakuin College, Japan)
10:00 a.m.	***Computer-assisted pronunciation learning of French /u/ and /y/ at the intermediate level*** **Viviane Ruellot (Western Michigan University)**
10:30 a.m.	Break
10:50 a.m.	*L2 speakers' impressions of the roles of accent and intelligibility after seven years in an ESL environment* Tracey Derwing (University of Alberta, Canada)
11:20 a.m.	***Acquisition of L2 phonology in advanced learners: Does instruction make a difference?*** **Anita Saalfeld (University of Nebraska-Omaha)**

11:50 a.m.	*Empowering students to develop their own voice: An achievable goal* Joanna Smith (Unitec, New Zealand)
12:45 p.m.	Lunch
1:45 p.m.	Posters
	The perception and production of English word-final alveo-palatals by Korean L1 learners of English Amanda Huensch (University of Illinois)
	Funny marks in French? C-cédilles and first semester learners Jessica Sturm (Purdue University)
	Vowel production abilities of Haitian-American children **Stacy Wallen (Georgia State University) & Robert Fox (Ohio State University)**
	Students' awareness of Spanish spirantization allophonic rule **Manuela Gonzalez-Bueno & Marcela Quintana-Lara (University of Kansas)**
	Intelligibility in second language communication: A two-way street Nicole Eustice (University of Iowa)
	Nonlanguage factors affecting nonnative undergraduate students' reaction to ITAs **Edna Lima (Iowa State University)**
	The acquisition of aspiration in Serbian EFL learners Biljana Cubrovic (University of Belgrade, Serbia)
	Using mobile technologies for synchronous CMC to develop L2 oral proficiency **Sarah Huffman (Iowa State University)**
3:00-4:30 p.m.	**Interactive panel discussion** ***What really affects and improves intelligibility?*** ***Moderated by John Levis***
5:45-8:30	Reception

The papers in the proceedings are divided into three sections that have broadly similar themes. There are five papers in the first section, all generally related to *Listener Reactions to Accented Speech*, including discussions of intelligibility and sociolinguistic reactions such as irritation. The second section also includes 5 papers and is titled *Pedagogical Approaches*. Finally, the last section includes six papers on the general theme of *Pronunciation Acquisition and Description*.

Listener Reactions to Accented Speech

The first section, *Listener Reactions to Accented Speech*, starts with the plenary address by **Murray Munro** of Simon Fraser University. "Intelligibility: Buzzword or buzzworthy?" examines the contention put forth in an opinion article in Applied Linguistics (Rajogapolan, 2010) that intelligibility is a politically charged concept that sounds good but is ultimately counterproductive in relation to research and teaching. Munro provides evidence that intelligibility has a long history both in and outside of linguistics, and that intelligibility is also an issue that has real-world implications in fields as diverse as aviation and public service encounters.

In "(Un)Intelligibility Tales," **Laura Hahn and Patricia Watts** discuss narratives provided by teachers and students about times they experienced loss of intelligibility. Most described exchanges involving NS-NNS interactions or those between two NSs. They analyze the exchanges according to the features of the speech exchange that caused the loss of intelligibility, how the confusion was repaired, and the role context played in remedying the loss.

Ghazi Algethami, John Ingram, and Thu Nguyen, in "The interlanguage speech intelligibility benefit: the case of Arabic-accented English," looked at the Interlanguage Speech Intelligibility Benefit, the hypothesized intelligibility advantage L2 listeners have over native listeners when they listen to speakers who share their native language. They found no significant difference between L1 Arabic speakers and native English speaking listeners when listening to Arabic accented speech. They interpret these findings in light of previous findings and suggest that phonetic features of L2 speech more strongly influence listeners' perceptions than native language.

In "Language and Nonlanguage Factors Affecting Nonnative Undergraduate Students' Reaction to ITAs," **Edna Lima** investigated the perceptions of international undergraduate students toward international teaching assistants (ITAs). In a matched-guise study, students enrolled in first-year composition classes watched a short video-taped lecture with three guises. Significant results were found for accent and speaker likeability.

Finally, **John Levis** provides a collection of themes that arose during a panel discussion of pronunciation experts from around the world in "Assessing Speech Intelligibility: Experts Listen to Two Students." The panel and audience listened to recordings of the English free speech and read speech of two students, a Spanish speaker and a Korean speaker. Panelists discussed the issues that affected the speech intelligibility of the two students. Four themes are examined: The role of listeners' experience, the features that most impacted intelligibility, the tendency to attribute loss of understanding to accent regardless of the true cause, and the problems with assessing using tasks that employ read speech.

Pedagogical Approaches

The second section of the proceedings includes a variety of papers that address pedagogical concerns. In "Upstream Destressing," **Wayne Dickerson** argues that destressing, a well-attested feature of English pronunciation after the sentence focus, also occurs in phrases in which two or

more strongly stressed syllables precede the sentence focus. He discusses how de-stressing the strongly stressed syllable immediately preceding the focus provides an environment in which the focused syllable can be heard more clearly and provides pedagogical suggestions for making use of this pattern in teaching focus.

Amanda Baker discusses how teachers' knowledge and beliefs develop and the ways in which those beliefs affect their classroom practice in "ESL Teachers and Pronunciation Pedagogy: Exploring the Development of Teachers' Cognitions and Classroom Practices." She examines the cognitions and practices of five experienced ESL teachers and finds that training in pronunciation pedagogy strongly affects not what they know but their confidence. In addition, their own language learning background, teaching experience and work with other teachers influence their approach to pronunciation.

In "An Integrated Approach to Pronunciation: Listening Comprehension and Intelligibility in Theory and Practice," **Marnie Reed** and **Christina Michaud** argue that the relationship between speaking and listening should be understood in terms of an auditory feedback loop, in which production facilitates perception. Four components are proposed for pronunciation teaching to accompany and reinforce other elements of language teaching: connected speech, suprasegmentals, inflectional morphology, and segmentals. Suggestions are given for what such an approach might look like in the classroom.

Veronica G. Sardegna, in "Pronunciation learning strategies that improve ESL learners' linking," explores how teaching pronunciation learning strategies improved linking within and across words. Teaching strategies led to initial and long-term improvement for all groups of students. The pronunciation learning strategies that were used are examined in detail and pedagogical implications are given for the use of strategies especially in regard to the teaching of linking.

In "Using Mobile Technologies for Synchronous CMC to Develop L2 Oral Proficiency," **Sarah Huffman** moves away from a narrow focus on pronunciation to explore how Computer Mediated Communication (CMC) using mobile technologies can help promote oral proficiency. She explores how multimedia mobile technologies have great promise in helping learners develop their L2 oral proficiency. The researcher also provides points to consider to make the implementation of mobile technologies more successful in teaching L2 oral language and suggests future research avenues.

Pronunciation Acquisition and Description

The last set of papers address descriptions of learner pronunciation and the way pronunciation is acquired. In "The Pronunciation of /s/ in Complex Onset and Coda Clusters in Somali-Accented English," **Ettien Koffi** describes how /s/ in complex onset and coda clusters patterns in the speech of Somali learners of English, a major immigrant population in Minnesota, a midwestern US state. He first describes how /s/ contributes to accentedness, and then explains the patterns using two phonological principles, the Sonority Sequencing Principle and the Coda Condition.

Anita Saalfeld analyzes the acquisition of Spanish stress in "Acquisition of L2 phonology in advanced learners: Does instruction make a difference?" She replicated previous research on the acquisition of Spanish stress with a group of intermediate learners in a Spanish phonetics course and a control group in another non-phonetics oriented Spanish course. Her results suggest that when students can self-enroll in courses focusing on pronunciation, those who need pronunciation instruction the most are more likely to avoid classes that include phonetics instruction.

In "Vowel Spaces in Bilingual Haitian American Kindergartners" **Stacey Wallen** and **Robert Fox** examine the vowels used by monolingual and bilingual Haitian American kindergartners. Because such bilingual speakers are often identified for help with communication disorders when they may simply have a communication difference, it is important to identify whether such help is necessary. They found that vowel production of Haitian American kindergartners differed from bilingual Haitians only in production of /o/, and that there were no significant differences in the vowels of bilingual Haitian American children and native English speaking controls.

In "The impact of computer assisted pronunciation training on the improvement of Vietnamese learner production of English syllable margins," **Lucy Pickering, Pam Pearson,** and **Rachel DaSilva** look at how computer assisted pronunciation training can be used to improve the pronunciation of Vietnamese learners of English at syllable margins. Thirteen intermediate Vietnamese students took part in eight 30-minute tutoring sessions where they used spectrograms of their own speech to compare their production to the prerecorded spectrograms. The methodology proposed in the study is put forth as a model for other pedagogical interventions based on the reactions of both teachers and students.

Manuela González-Bueno and **Marcela Quintana-Lara** examine the extent to which L2 Spanish learners who vary in language proficiency are aware of how Spanish voiced stops become fricatives in particular environments in their paper"Students' Awareness of Spanish Spirantization Allophonic Rules." The results indicate that proficiency is strongly related to awareness, with intermediate learners beginning to show awareness, and that the development of awareness starts earlier for velars and dentals than for bilabials. In production, the sounds not found in English (velar and bilabial fricatives) were produced more frequently than dental fricatives (a sound found in English).

In "Computer Assisted Pronunciation Learning of French /u/ & /y/," **Viviane Ruellot** explores how visual feedback improves the pronunciation of French /u/ and /y/ in the intermediate-level adult learners of French. Because /y/ is not an English sound, learners will not acquire it without sufficient exposure. Even though visual reinforcement has been successful for other pronunciation features, this was not the case for these vowels.

References

Rajagopalan, K. (2010). The Soft Ideological Underbelly of the Notion of Intelligibility in Discussions about 'World Englishes.' *Applied Linguistics, 31*, 465-470.

Munro, M. J. (2011). Intelligibility: Buzzword or buzzworthy?. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (pp.7-16),Ames, IA: Iowa State University.

INTELLIGIBILITY: BUZZWORD OR BUZZWORTHY?

Murray J. Munro, Simon Fraser University

Intelligibility is a much-touted concept in current research on second-language pronunciation teaching and in discussions of World Englishes. Some recent commentary has even asserted that the term has become a meaningless buzzword. However, interest in this very central aspect of language communication has a long history. Over a century ago, Henry Sweet's (1900) guide to practical language study made numerous references to it, and throughout the 20th century it has been of interest not only to language teachers, but also to a wide range of researchers and practitioners in the speech and communication sciences. Intelligibility is a fundamental requirement in human interaction, while the costs of *unintelligibility* range from minor inconvenience to matters of life or death. Although a focus on intelligibility has important repercussions for language teaching, social interaction, identity, and even human rights, defining the concept and determining its underpinnings have posed major challenges for researchers. In fact, many aspects of the notion remain poorly understood. This discussion examines the origins and significance of the construct, and identifies directions for future research on intelligibility that will help to establish the most effective ways of achieving it.

THE INTELLIGIBILITY CONSTRUCT

Of all the ideas embraced by contemporary pronunciation specialists, the notion of intelligibility as a sound teaching principle seems uncontroversial. In fact, it almost enjoys the status of a hallowed pursuit. And why shouldn't it? If, by intelligibility, we mean the extent to which utterances are understandable to a speaker's audience, it is difficult to conceive of many realistic circumstances in which *any* speaker would not want it. Language teachers therefore take it as self-evident that learners want and need to develop intelligible speech patterns, a fact reflected in the growing interest in how that end can be achieved.

Given the wide consensus on these matters, it seems astonishing to see intelligibility identified as a suspicious, politically-charged concept that cannot be justified in applied linguistics. Yet Rajagopalan (2010) does just that in asserting that intelligibility is merely a buzzword (p. 465) on a par with value-laden adjectives like *beautiful*, *ugly*, and *primitive*, which, according to him, "we have long learned to regard with suspicion" (p. 465). It is possible that his assertions arise from a misunderstanding of how the term *intelligibility* is applied by those concerned with fostering effective communication skills in ESL speakers. But whatever the explanation, his claims are patently false.

First, let's consider the statement that intelligibility is a buzzword – voguish and soon to be forgotten. Outside of language pedagogy, that is certainly not true. Telephone companies, sound engineers, architects, speech-language professionals, air traffic controllers and many others have been deeply concerned with intelligibility for many decades. The result is an accumulation of

thousands of research papers in technically-oriented publications like the *Journal of the Acoustical Society of America*, many dating back to the early 20th century. If that concern had never existed, our daily lives would be much the poorer – telephones would not transmit speech optimally; school classrooms would have poor acoustics, such that children would have difficulty understanding their teachers; and patients with speech disorders would have no hope of communicating effectively with family members.

It is not known when intelligibility was first discussed in connection with language teaching. However, the term goes back at least as far as Henry Sweet's (1900) book on practical language study. Sweet saw intelligibility as a guiding principle in the teaching of pronunciation, which, for him, was foundational in L2 learning. A half-century later, Abercrombie (1949) published his well-known article, articulating for perhaps the first time the dictum that second language learners should aim for a "comfortably intelligible" pronunciation rather than for a native-like accent. Still later in the 20th century, Gimson (1962) published his popular volume on English pronunciation. Revised by Cruttenden and now in its eighth edition (Gimson, 2008), it refers extensively to intelligibility, emphasizing that learners need not sound like native speakers. Gimson goes even so far as to suggest how L2 users can be understood by substituting particular consonants for ones they are unable to produce. Meanwhile, over the past 40 years, intelligibility has taken its place in a wide range of pedagogical materials (e.g., Rogerson & Gilbert, 1990), as well as in research agendas relating to World Englishes (Smith & Bisazza, 1982; Nelson, 2008) and L2 pedagogy (Munro & Derwing, 1995; in press). Intelligibility has been an active area of discussion and study for a very long time. Nor is there any reason to believe that concern about this topic will diminish in the near future. In short, far from being a buzzword, *intelligibility* is a well-established construct with a firm foundation in empirical and pedagogical traditions.

Second, let's compare *intelligible* with the adjectives *beautiful*, *ugly*, and *primitive* – all highly subjective terms. In fact, determining the intelligibility of speech does necessitate a response from an interlocutor, and the concept is evaluative in the sense that the intelligible is preferred over the unintelligible. But the difference lies in the reason for that preference. We prefer beautiful things and disprefer ugly ones as a matter of personal choice; ultimately, not much hinges on whether our preferences are honoured. As far as I know, no one has ever died from viewing something and judging it ugly or primitive[1]. But a loss of intelligibility has at times resulted in human tragedy. Over the past several decades, the NASA aviation reporting system has documented thousands of cases of communication breakdowns between pilots and air traffic controllers, many the result of unintelligibility. In some cases, a specific phonetic ambiguity was directly tied to tragedy. Cushing (1995) provides the following example:

> ATC cleared the aircraft to descend "two four zero zero." The pilot read back the clearance as, "OK. Four zero zero." The aircraft then descended to 400 feet (122 meters) rather than what the controller had meant, which was 2,400 feet (732 meters). (p. 3)

Tragically, as a result of this stress-related misinterpretation of *two* as *to*, four crew died. Even more troubling is that, according to McMillan (1998), this specific confusion is common in

[1] Here I am excluding from consideration any of those unfortunates in Greek mythology who looked upon the grotesque head of Medusa and were turned to stone.

aviation. Given that sobering fact, I leave it to you – the next time you are in an airplane – to ponder how comparable *intelligible* is to *beautiful* or *primitive*.

In second language (L2) classrooms, we don't expect to encounter life and death situations. But the ramifications of intelligibility loss for L2 learners are very real and very serious. Here we need only consider the plight of immigrant language learners who find that people don't understand them. Apart from the personal frustration they may feel, communication difficulties can damage educational and career opportunities. They can also lead to negative social evaluation, such that others avoid interactions because conversation seems a difficult chore. And that can lead to isolation from the host community and lost opportunities to use the L2. Denying that these serious consequences exist is not simply a mistaken view; it is a reckless and irresponsible position that dismisses the communicative needs of L2 learners.

EMPIRICAL APPROACHES TO INTELLIGIBILITY

In our research program, my colleague Tracey Derwing and I have developed a tripartite perspective on the study of L2 pronunciation (Derwing & Munro, 1997, 2005; Munro & Derwing, 1995). We distinguish among *accentedness* – how different someone's speech seems (often from the listener's variety), *comprehensibility* – the listener's experience of how difficult the speech is to understand, and *intelligibility* – how much of the speech is actually understood by interlocutors. Of these, *accentedness* is the least relevant to communication because listeners quickly adjust to many divergences from their own speech patterns; many L2 speakers have strong accents but are perfectly intelligible to their interlocutors. *Intelligibility* and *comprehensibility*, however, are more closely tied to communicative success, and therefore merit careful attention. One particularly important empirical finding is that it is possible for learners to become more intelligible or comprehensible through instruction with no noticeable change in accentedness (Derwing, Munro & Wiebe, 1998). Furthermore, when aspects of an accent change, there is no guarantee that benefits in intelligibility or comprehensibility will necessarily accrue. Here I identify some approaches to studying these dimensions and offer suggestions for further work.

Focusing on phonetics

Recent empirical work often aims at identifying phonetic properties of L2 speech that reduce intelligibility and comprehensibility. This focus represents an important shift, in that much previous research addressed the quite different question of whether teaching or lab training could improve perception or production of particular speech sounds. The problem with the latter orientation is its failure to distinguish between improvement that genuinely makes a difference for communication and improvement that is less consequential or mainly cosmetic. When we recognize that pronunciation instruction cannot possibly receive unlimited attention in the classroom, it becomes clear that teachers need to set priorities. Some research provides us with useful guidance on how to do this. It is now largely uncontroversial, for instance, that certain L2 prosodic difficulties undermine intelligibility, as demonstrated by Hahn's (2004) examination of primary (nuclear) stress. In that study, listeners' processing of utterances with and without primary stress errors was compared. A different approach was used by Tajima, Port, and Dalby (1997), who improved intelligibility of recorded L2 speech by digitally manipulating rhythmic patterns. And a third technique (Derwing et al., 1998) revealed improved comprehensibility in

narratives as a result of global prosodic instruction. Thus prosodic difficulties are not only detrimental to communication; they can be overcome.

With respect to the teaching of segmentals, pronunciation specialists generally accept that some vowel and consonant distinctions deserve higher priority than others. Statistically, low functional load (FL) contrasts in English, such as /θ/ – /f/ and /ð/ – /d/, distinguish relatively few common word pairs in English, so that confusions of these segments might be predicted to have minimal communicative impact. High FL distinctions, like /l/ – /n/ and /l/ – /ɹ/, on the other hand, should be important to maintain. These predictions are theoretically-driven (Brown, 1991; Levis & Cortes, 2008) and require empirical verification. So far, limited testing (Munro and Derwing, 2006) supports the importance of FL for comprehensibility, suggesting also that high functional load substitutions have cumulative effects on listeners, whereas low FL substitutions do not.

The promising techniques described above have already yielded useful results for the L2 classroom; further work on such matters as functional load is likely to provide more detailed evidence about the phonetic sources of unintelligibility.

Broadening our view

Despite its many merits, a purely phonetic approach to intelligibility is insufficient. In the first place, some intelligibility breakdowns lack a straightforward phonetic explanation. McMillan (1998), for instance, reports another aviation accident in which an air traffic controller instructed a pilot to "Take taxiway right" (p. 44). Instead, the pilot heard, "You can backtrack if you like." Although the two sentences are segmentally and rhythmically similar, it is impossible to establish how this transformation occurred in the mind of the pilot. Moreover, we cannot derive any clear lesson from this example to prevent a comparable error in the future.

A similar problem arose in an analysis by Derwing and Munro (1997). When describing a standard picture story, a Cantonese speaker produced an utterance that has proved unintelligible on first encounter to virtually everyone who has ever heard it. To our ears, the speaker seemed to say "one man dry cuckold," and it was only after multiple hearings that we determined the intended words: "One man drive car, go…" Here grammar and pronunciation, including a high pitched monotone intonation and staccato rhythm, conspired to cause unintelligibility. The "cuckold sentence" and the aviation example illustrate how listeners' perceptual systems attempt to find meaningful units within a stream of speech that does not conform well to familiar patterns. Reliance on top-down processes may lead to mondegreens – misinterpretations because of multiple ambiguities in the speech that interact in complex ways.

It is easy to forget that the causes of unintelligibility are not restricted to the speech itself. Much daily communication takes place under non-ideal listening conditions. In fact, many immigrants work in noisy environments – in food service locations, hospitals, and factories. Like all of us, they encounter noise and distortion during telephone conversations and public address announcements at airports. Even proficient speakers may experience difficulties in such circumstances. However, for L2 learners, the problems are magnified. On the one hand, L2 listeners experience a greater loss of comprehension than native listeners in noisy and reverberant conditions (Takata & Nábelek, 1990). On the other hand, when noise is present, L2

speech sometimes undergoes a greater decrement in intelligibility to proficient listeners than does L1 speech. In Munro (1998), for instance, I added cafeteria noise to native and Mandarin-accented English utterances, and presented them to listeners (along with noise-free speech). In the noisy condition, the intelligibility of both L1 and L2 speech was compromised; however, many L2 utterances showed a much more dramatic reduction in intelligibility than did the L1 speech. The reduction also varied considerably across speakers. While some showed moderate effects, one L2 speaker was about 85% intelligible in quiet, but dropped to less than 10% in noise. Thus, some L2 speech is more noise-resistant than other speech.

Native speakers reflexively adjust intensity, pitch, and rhythmic characteristics when speaking in noise – the *Lombard Effect* (Van Summers et al., 1988). However, little is known about the effectiveness of these and other possible adjustments on foreign-accented speech. An intriguing question is whether L2 speakers can learn techniques for enhancing their speech intelligibility in noise. Currently, pronunciation instruction is typically carried out in relatively quiet classrooms or language labs. Experience with a wider range of speaking and listening conditions may turn out to be a valuable complement to regular classroom work.

The flip side of non-ideal listening conditions is non-optimal speaking techniques, another concern that has little to do with phonetics. It encompasses such matters as poor vocal projection, excessive glottal fry (very low-pitched speech of weak intensity), covering one's mouth while speaking, ineffective pausing, and a host of other behaviors. Of course, these habits can be detrimental for all speakers; however, when an L2 accent is simultaneously present, it may be mistakenly identified as the chief source of unintelligibility. Such *accent scapegoating* may account for some of the negative judgments of International Teaching Assistants (ITAs) that have been reported for a number of years. In such cases, the solution may have little to do with accent; it may lie in teaching ITAs better classroom speaking skills, heightening their awareness of cultural expectations, and encouraging them to monitor their audience for comprehension.

Attending to the listener

Recent research examining listener factors in intelligibility serves as a valuable reminder of the two-way nature of oral interaction (Kennedy & Trofimovich; Zielinski, 2008): successful communication depends on the abilities and efforts of both speaker and listener. Listeners with certain experience, background, and perhaps aptitude may be more successful than others at comprehending L2 speech. For example, familiarity with a particular L2 accent may aid in comprehension of speakers with that accent (Gass & Varonis, 1984). Also, listeners from a particular L1 background may have an advantage in understanding English L2 speakers from a shared background. Although evidence supporting both predications exists, the size of the effects has been small and inconsistent (Major, Fitzmaurice, Bunta, & Balasumbramanian, 2002; Munro, Derwing, & Morton, 2006). A possible reason is insufficient linguistic information available in the speech. The "cuckold sentence" mentioned earlier was no better understood by listeners who had the same L1 accent as the speaker than by anyone else. Perhaps the speech lacked the phonetic and grammatical characteristics that any listener would need in order to understand it readily. If so, there is no reason why a listener from a shared L1 background would have any advantage.

While the effect of experience on listeners appears to be small, the similarities in responses to L2 speech between listeners are remarkable – even when native and non-native listeners are compared. For example, Derwing and Munro (under review) compared comprehensibility ratings from native English listeners and a group of high proficiency non-native listeners from mixed L1 backgrounds. The correlation between the groups' mean scores (Pearson *r*) was .94, indicating that the two groups strongly agreed on which speech samples were hard to understand and which were easy. While further work on the effects of experience and L1 background is still needed, researchers should consider exploring other listener factors that may prove more important.

One potentially influential variable is age. Burda, Scherz, Hageman, and Edwards (2003) have used our comprehensibility measure to evaluate perception by geriatric listeners with age-typical hearing acuity. In general older listeners have greater difficulty than younger listeners at understanding an accent other than their own. While the effect may be partly due to hearing loss, central processes in speech perception deteriorate with age. As a result, geriatric listeners have a harder time processing speech in general, and a novel accent may become particularly challenging.

The role of aging in speech perception deserves attention in immigrant-receiving countries like Canada. First, increasing numbers of immigrants work in health care and elder care. Second, linguistic diversity is increasing among the elderly as older immigrants enter the country and the existing immigrant population ages. Therefore, among geriatric patients and their caregivers, the likelihood of interactions involving different accents is increasing. This raises important issues for the training of health care workers. In particular, they need sensitivity to linguistic issues – especially intelligibility – and strategies for effective interactions. Research on speech in medical contexts has already taken some interesting directions, as in Radonovich, Yanke, Cheng, and Bender (2010), who found substantial declines in intelligibility due to the use of respirators and masks frequently worn by medical staff. Expansion of this line of work could prove interesting and practically useful.

Another aspect of the bidirectional nature of interactions is that when communication fails, it is incorrect to automatically assume that the speaker is at fault. It is well established that attitudinal and motivational factors affect how listeners respond to L2 speakers, such that misunderstanding may have much more to do with a listener's investment in an interaction than with aspects of the speech itself. For instance, prejudice against particular accents or against immigrants in general can distort listeners' perceptions during interactions and result in accent discrimination (Munro, 2003). Regrettably, the "accent reduction" industry often exploits immigrants' insecurities about their accents as a way of marketing their dubious and expensive services. But prejudice is not based on rational thinking, and attempting to adjust one's accent to placate those who discriminate is likely to be a fruitless endeavor. In some cases, however, native speakers' difficulties in interactions with L2 speakers are due to fear and inexperience. In such cases, evidence indicates that native speakers can change their outlook on L2 speech to become more receptive to such interactions (Derwing, Rossiter, & Munro, 2002). Further work on how intelligibility can improve through listener training has potential benefits for many contexts, including workplaces in which immigrants work closely with others from diverse L1 backgrounds. In immigrant-receiving countries we can expect such an approach to become more and more important in the future.

CONCLUSION

Intelligibility is the single most important aspect of all communication. If there is no intelligibility, communication has failed. In language pedagogy this is not a new idea; nor is the current interest in intelligibility a passing fad. Rather it is an empirically sound concept that will provide a basis for a wide range of pedagogically-oriented research in the future. While some such work will continue to focus on phonetic issues, more consideration needs to be given to the effects of speaking and listening conditions on L2 speech and to speaker behaviors that facilitate comprehension. We also need to expand our investigations to a wider range of listener factors such as aging. A final, potentially fruitful line of work is the development of ways to improve people's receptivity to different patterns of speech.

Many are familiar with Jonathan Swift's *Gulliver's Travels*, written around 1726. Swift was a social critic with a particular contempt for navel gazing. On one journey, Gulliver visits Laputa[2], an island that floats in the air and is populated by an intelligentsia that has control over the country of Balnibarbi below. Preoccupied with abstraction, the academics develop theoretically-based ways of doing almost everything that are adopted by the Balnibarbians. When houses are built, for instance, right angles are not allowed. To fit clothing, a tailor must not use a tape measure. The result is a country in shambles. The population is dirty and hungry, and walks around in rags.

L2 researchers would do well to heed Swift's lampooning of academia. If they are to be taken seriously by language teachers and their students, they must not lose sight of the accumulation of knowledge pointing to the centrality of intelligibility in communication. Among applied linguists there is no shortage of ideas about what is and is not important. But bad ideas – especially those motivated by overweening, abstract argumentation rather than practical realities – must not be allowed to trump learners' needs. In particular, we do not need to debate the issue of whether intelligibility is important. Rather, we need to carry on with our work on how we can apply this concept in the most effective ways.

ACKNOWLEDGMENTS

The ideas in this manuscript are the product of a long-term collaboration with my colleague Tracey Derwing, to whom I owe much gratitude. I also thank our co-authors, students, and RAs for their invaluable contributions to our research, as well as our main funder, the Social Sciences and Humanities Research Council of Canada.

ABOUT THE AUTHOR

Murray Munro is a Professor of Linguistics and Director of TESL at Simon Fraser University. A former ESL teacher he is a specialist in applied phonetics, and has carried out research on such topics as age and accent, the role of accents in communication, L2 acquisition of vowels, and the social evaluation of L2 speakers. His postdoctoral work was completed in James Flege's research lab. Since then he has collaborated extensively with Tracey Derwing. In 2007 he co-edited

[2] It is probably no coincidence that *la puta* in Spanish translates to *the prostitute.*

Language Experience in Second Language Speech Learning (Benjamins), and, with John Levis, he is currently an area editor (Phonetics and Phonology) of the 10-volume *Encyclopedia of Applied Linguistics* (Wiley Blackwell).

REFERENCES

Abercrombie, D. (1949). Teaching pronunciation. *English Language Teaching, 3*, 113– 122.

Brown, A. (1991). Functional load and the teaching of pronunciation. In A. Brown (Ed.), *Teaching English Pronunciation: A Book of Readings* (pp. 211–224). London: Routledge.

Burda, A. N., Scherz, J. A., Hageman, C. F., & Edwards, H. T. (2003). Age and understanding of speakers with Spanish or Taiwanese accents. *Perceptual and Motor Skills, 97*, 11–20.

Cushing, S. (1995). Pilot – air traffic control communications: It's not (only) what you say, it's how you say it. *Flight Safety Digest, 14 (7)*, 1–10.

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. (2009). Comprehensibility as a factor in listener interaction preferences in the workplace. *Canadian Modern Language Review, 66*, 181-202.

Derwing, T. M., & Munro, M.J. (Under review). The Development of L2 Oral Language Skills in Two L1 Groups: A Seven-Year Study.

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.

Derwing, T. M., Rossiter, M. J., & Munro (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingual and Multicultural Development*, *23*, 245–259.

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

Gimson, A. C. (1962). *An Introduction to the Pronunciation of English*. London: Edward Arnold.

Gimson, A. C. (2008). *Gimson's Pronunciation of English, 8th Edition*. Revised by A. Cruttenden. London: Hodder Education.

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

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review, 64*, 459-489.

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.

Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL Quarterly, 36*, 173–190.

McMillan, D. (1998). *Miscommunications in Air Traffic Control*. Unpublished MA Thesis, Queensland University of Technology.

Munro, M. J. (1998).The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition*, *20*, 139-154.

Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal*, 20(2), 38-51.

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

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. (In press). The foundations of accent and intelligibility in pronunciation research. *Language Teaching*.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*, 111–131.

Nelson, C. L. (2008). Intelligibility since 1969. *World Englishes, 27*, 297–308.

Radonovich, L. J., Yanke, R., Cheng, J., & Bender B. (2010). Diminished speech intelligibility associated with certain types of respirators worn by healthcare workers. *Journal of Occupational and Environmental Hygiene*, *7*, 63-70.

Rajagopalan, K. (2010). The Soft Ideological Underbelly of the Notion of Intelligibility in Discussions about 'World Englishes.' *Applied Linguistics*, *31*, 465-470.

Rogerson, P., & Gilbert, J. (1990). *Speaking Clearly: Pronunciation and Listening Comprehension for Learners of English*. Cambridge: Cambridge University Press.

Smith, L. E., & Bisazza, J. A. (1982). The comprehensibility of three varieties of English for college students in seven countries. *Language Learning*, *32*, 259–269.

Sweet, H. (1900). *The Practical Study of Languages*. New York: Henry Holt and Company.

Tajima, K., Port, R., & Dalby, J. (1997). Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics, 25*, 1–24.

Takata, Y. & Nábelek, A. K. (1990). English consonant recognition in noise and in reverberation by Japanese and American listeners. *Journal of the Acoustical Society of America, 88*, 663–666.

Van Summers, W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., & Stokes, M. A. (1988). Effects of noise on speech production: Acoustic and perceptual analyses. *Journal of the Acoustical Society of America, 84*, 917-928.

Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. *System, 36*, 69–84.

Hahn, L. & Watts, P. (2011). (Un)Intelligibility Tales. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 17-29),Ames, IA: Iowa State University.

(UN)INTELLIGIBILITY TALES

Laura Hahn, University of Illinois at Urbana-Champaign
Patricia Watts, University of Illinois at Urbana-Champaign

We collected forty-two narratives describing communicative events in English that involved a misunderstanding attributed to pronunciation. The majority of the narratives described exchanges involving either a native speaker and a non-native speaker, or two native speakers. In this paper we report on the kinds of phonological features involved in the miscues in the exchanges, the kinds of repair strategies used, and whether and how participants used context to help them interpret the utterance. We also discuss elements in the narratives that describe the analytical and emotional reactions of participants in these exchanges. Finally, we reflect on the use of stories of (un)intelligibility for pedagogical and research purposes.

INTRODUCTION

We begin our paper with a story from our data.

> *It was my first ever visit to Africa. I had arrived at Kano, Nigeria, the day before, and here I was at the Ahmadu Bello University, Zaria, on Sunday morning, all by myself before my wife and family were due in 6 weeks' time. I went out to survey the garden of the accommodation that had been allocated to me, and was immediately impressed by the colour of the flowers and birds, and so I went back in to fetch my camera. When I came out again with my camera, there was a young Nigerian lady dressed up in her Sunday finery, accompanied by 3 young children. She looked at me and smiled and said, "You snuff me?" I was shocked and bewildered; what on earth was she suggesting? The rising tone clearly meant to me a request or an invitation, but to snuff her!? In front of 3 little children at that? No, she didn't say "sniff," but that itself would have been just too embarrassing and out of the question. I wouldn't dream of trying to appreciate her scent! But "snuff" sounded even worse; what was she hoping I would do in front of the children? Help! I've only just arrived and here was a young lady making a proposition to me!*
>
> *She could obviously tell that I was bewildered, and so she helped me out by pointing to my camera and smiling. Then it struck me! She was asking me to take a photograph of her in her Sunday best; she wanted me to take a snap of her. So overcome with relief, I willing did so, with the little children in tow.*
>
> *"To snap a person" means to take a photograph of them. Hausa has no /p/; so [f] substitutes in Hausa English. Hausa articulates its /a/ as [ʌ], a mid central vowel. Hence, British English "snap" is produced as "snuff" by a Hausa speaker with limited English. Phew!*

Stories such as this one may have productive uses beyond their role as anecdotes: they may help untangle some of the complexities involved in understanding how intelligibility works.

Stories of unintelligibility reveal several parts of the miscommunication. First, there is the original unintelligible utterance, in this case "snap." Second, there is the recognition of the problem; in this story, the Nigerian lady shows her recognition by pointing to the camera. And third, stories can allow us to examine any repair that took place; here, the Nigerian lady did not change her pronunciation, although the meaning was resolved. Stories may also capture emotional experiences (frustration, aha moments, etc.) and other relevant details, e.g. "I was dead tired that day." In the story above, we see the storyteller's bewilderment and eventually relief. We also see that he tries to use context to identify (and discard) possible interpretations of the word in question, and that he eventually uses academic background in phonology in analyzing the event.

As this story reveals, both the speaker's and listener's contributions and perspectives are critical to the "intelligibility cocktail" (Zielinski, 2006) – and we posit that stories provide a rich, qualitative understanding of the ingredients.

RESEARCH QUESTIONS AND METHODOLOGY

We examined forty-two first-person accounts of "unintelligibility tales" from native and non-native speakers, describing NNS – NS and NS – NS exchanges. Our broad research question was: What patterns emerge related to the elements of unintelligible utterances in NNS – NS and NS – NS exchanges? More specifically, we were interested in:

> The kinds of phonological features involved.
>
> The kinds of repair strategies used.
>
> Whether and how participants used context to help them interpret the utterance.

We collected eight stories from international graduate students at the University of Illinois who were in ESL classes; the remainder of the stories came from native or near-native speakers -- our language teacher colleagues, our personal acquaintances, and ourselves. Thirty-two of these stories were written narratives; the remaining eleven were told (or happened) to us, and we transcribed the oral accounts. The prompt for the written narrative is in Appendix A. Table 1 delineates the types of stories we examined, with examples (the misinterpretation is first, followed by the intended meaning).

Table 1. *Categories and Examples of Misunderstandings.*

		NS mishears		NNS mishears
NS utterance	11	gorgeous guy, gorgeous sky She's engaged, She's in Ames	6	mammals, Mormons/ marmots condom, condo
NNS utterance	22	meth camp, math camp dog, duck	3	TOEFL, tofu grammar, grandma

FINDINGS AND ANALYSIS

Phonological Features

Bond (1999, p. 61) states, "The most revealing way of describing complex misperceptions is to consider how the phonological shape of the misperception matched or failed to match the target utterance." To this end, we analyzed the phonological features that contributed to misunderstandings in the words and phrases in our data. Because the numbers of cases of NNSs mishearing NSs and NNSs are relatively low, we focused our analysis instead on cases of NSs mishearing either NNSs or other NSs.

Our classifications were based on Zielinski (2006), in which she summarizes research on segmental and suprasegmental features that can mislead listeners in their attempts to identify spoken utterances. We tagged each utterance in our data according to whether the following features played a role in the miscommunication: vowels, consonants (word-initial, medial, and final positions), occurrence in weak or stressed syllable, syllable insertion or deletion, word or compound stress, and word boundary confusions.

In 65% of the misunderstood utterances in our data, there was more than one phonological feature involved. For example, when a native speaker interpreted a Brazilian speaker's "Bed time?" as "Bad time?" both vowel substitutions (/æ/ for /ɛ/) and misplaced compound stress contributed to the misunderstanding. When a native speaker's "in-house test" was heard by another native speaker as "enhanced test," both vowel (/aw/ - /æ/) and consonant (/s/ - /n(t)s/) features were at play.

For our analysis, we have to rely on the storyteller's version of how the utterances were pronounced and heard. In some cases, we can be fairly sure that a pronunciation problem caused the misunderstanding. In other cases, it seems more likely that the misunderstanding was due to a mishearing (perhaps due to lack of schema, or lack of attention, on the part of the listener). For example, it is improbable that "She's in Ames" would be pronounced as "She's engaged" by a native speaker.

Table 2 summarizes the analysis of phonological features involved in misunderstandings between NNS speakers and NS hearers. Again, the misinterpretation is first, followed by the intended meaning.

Table 2. *Native Speakers Mishearing Non-native Speakers*.

Feature	Examples		Frequency
Vowels	meth dog patience	math duck passion	(13/22)
Consonants	one in Ben Folds	one inch Penfolds	(16/22) Initial: 4 Medial: 3 Final: 9
Stress	impotent your surname	important user name	(7/22)
Syllable insertion or deletion	reject two in	register twin	(4/22) Insertion: 1 Deletion: 3
Strong or weak syllables	pork When is Atsuko due? passion	fork What does Atsuko do? patience	Strong: 14/22 Weak: 1/22 Both: 2/22
Word boundaries	your surname two in bad room	user name twin bedroom	(4/22)

In our data, the most commonly problematic vowel sound was /æ/. We note that many non-native speakers of English struggle to produce this sound. We also note that its production is variable and changing among native speakers (e.g. De Decker & Nycz, 2005). These two factors may account for the challenges involved in utterances containing this sound.

Among consonants, we found that stops were more frequently problematic than consonants with other manners of articulation. In addition, consonant deletion was a common cause of misunderstanding. For example, the deletion of final /tʃ/ resulted in "one inch" being interpreted as "one in." In this data set we had no cases of consonant insertion causing a problem.

In classifying words according to whether the problem occurred in a strong or weak syllable, one challenge arose: When a NNS stresses a syllable that should have been unstressed (e.g. "your surname" for "user name"), how should that be classified? We excluded two such cases from the counts in this category.

Table 3 summarizes the analysis of phonological features involved in misunderstandings between NS speakers and NS hearers.

Table 3. *Native Speakers Mishearing Native Speakers.*

Feature	Examples		Frequency
Vowels	Texas	taxes	(3/12)
Consonants	gorgeous guy you	gorgeous sky you'll	(8/12) Initial: 5 Medial: 6 Final: 2
Stress	inválid	invalid	(1/12)
Syllable insertion or deletion			(0/12)
Strong or weak syllables	in Ames Alan Prasanta	engaged Al and Prasanta	Strong: 6/12 Weak: 8/12
Word boundaries	Super salad	Soup or salad?	(4/12)

It is not surprising that no miscommunications between native speakers involved syllable insertion/deletion. The only word stress miscommunication occurred when a NS misread aloud "inválid" for "ínvalid." Proportionately, vowel problems were substantially lower for these encounters compared to our NS – NNS data (though we note that /æ/ was at play in two out of the three cases). Furthermore, word boundary problems were more salient for NS – NS exchanges than in NS – NNS. Interestingly, in looking at all 42 cases, there was only one instance of word boundaries causing a mishearing for a NNS ("To Kill a Mockingbird" was heard as "Tequila Mockingbird"), while word boundary issues affected six exchanges involving NS listeners.

Repair Strategies

In addition to investigating the nature of unintelligible utterances, one of our primary goals was to gain insight into how speakers attempted to repair communication breakdowns. For the twenty-two NNS-NS cases with NNS mispronunciations, active repair occurred in nineteen of these cases. Speakers used the following strategies: repetition (63.2%), providing additional information (36.8%), using non-verbal communication, such as pointing (15.8%), one case of spelling (.05%), and one case of paraphrasing (.05%).[i]

These findings differ markedly from Derwing & Rossiter (2002). In their study, subjects self-reported using paraphrase as their preferred repair strategy (56%) and repetition was mentioned as the second favorite strategy (28%). One possible explanation for these differences may be the nature of data collection. In Derwing & Rossiter, subjects were asked to state their preferred repair strategy, while in our study actual use of the strategy is reported. It is possible that Derwing & Rossiter's subjects could have under-reported a predilection for repetition or over-reported use of paraphrasing. Nevertheless, the almost complete lack of paraphrase by non-native speakers in our study is perplexing. Additional data about the speakers' level of English and explicit strategy training may have provided valuable insight into the dearth of paraphrasing, but without this information, it is hard to speculate. Additionally, the

exceptionally high percentage of repetition should be tempered by the fact that 33.3% of the speakers using repetition utilized other strategies in conjunction with repetition to clarify the unintelligible utterance. One noteworthy illustration of a NNS using a combination of strategies was related in the account of a NNS shopping for a watch. The watch she wanted had a duck on its face. Her request for the watch with *duck* was heard as *watch with dog*. After repeating *duck* unsuccessfully, the speaker wisely shifted her strategy, providing additional information (*"the duck with wings, not the dog that barks")* and pointed to the desired watch in the display case. The use of these strategies in combination ultimately led to a successful negotiation of meaning.

The twelve cases of NS-NS miscommunication provide interesting data for comparison. Repair occurred openly in seven of these cases. Here, too, repetition was the preferred strategy, appearing in 100% of the cases. In fact, only one other strategy appeared at all (providing additional information), and it was used only one time and in conjunction with repetition, the native speaker repeating *taxes*—which was misheard as *Texas* and adding the comment "*ya know, IRS*" for further clarification.

In some ways it is not surprising that repetition was the most common strategy for both NS and NNS. It is relatively simple and efficient to try repeating an utterance, especially if the speaker assumes that the second utterance attempt will help overcome potential language barriers, such as noise, listener inattention, or an utterance that wasn't articulated loudly or clearly enough.

While both NNS and NS relied heavily on repetition, the two groups' repair sequences were distinctly different. Notably, NSs use of repetition was effective on the first attempt 100% of the time. Thus, repetition sufficed in resolving the misunderstanding. For NNS, the success rate dropped to 33.3% and the NNSs in this study often were unable to channel any strategies beyond repetition. The steep decline in effectiveness of using repetition as a repair strategy can likely be attributed to NNSs phonological shortcomings and lack of ability to produce an utterance with native-like intelligibility even on demand.

Relevance of Context

A third question we investigated was the role context played in decoding unintelligible utterances. Analysis of contextual information in all forty-two stories yielded four classification categories. Table 4 displays data and examples related to context. In the final column, the misinterpretation is followed by the intended meaning.

Table 4: *Relevance of Context*

Role of Context	**Percentage**	**Context**	**Utterance**
Added to the confusion	34.9	Tailor shop	One in, One inch
Not relevant	18.6	Asking a stranger for directions	Unintelligible, McDonald's
Should have been useful but wasn't	32.5	Discussion about housing	Condom, Condo
Useful	14	Computer help line	Your surname, User name

The high percentage of cases in which context contributed to the confusion is noteworthy. In these stories, multiple interpretations of the unintelligible utterance fit with the context. Such was the case in a NS-NNS conversation about a student studying early childhood education. Here, the NS listener pondered whether she has heard "She must have a lot of passion" or "She must have a lot of patience, " both of which are contextually appropriate utterances. Despite the fact that context is not always available or helpful, we still believe that using context to disambiguate meaning is, overall, a productive strategy that listeners should keep in their repertoire of strategies when trying to decode an unintelligible utterance.

DISCUSSION

Zielinski (2006, p. 25) reminds us that "listeners draw on knowledge including context and syntactic and lexical knowledge to identify words in connected speech." In the stories we collected, we found instances of these factors playing (or not playing) a role. For example, in both of the "soup or salad" stories, the tellers revealed that unfamiliarity with the script of a sit-down restaurant contributed to interpretation of that phrase as "super salad." In the nearly all of our examples, what the hearer misheard still matched the part of speech of the original utterance. We also found a number of cases where the original word (or use of the word) was unfamiliar to the hearer, resulting in mishearing, for example, "snuff" for "snap," and "Daytown" for "Baytown." Finally, we observed that short answers and elliptical phrases yield less redundancy than their full forms, perhaps leading to ambiguity. An example is this exchange: "What are you doing this weekend?" "Taxes." "You're going to Texas?" Here, had the second speaker said, "I'm working on my taxes," the miscommunication would probably not have occurred.

Indeed, our data confirm the complexities involved in miscommunications. They also elucidate the substantial (yet erratic) role of context. However, we seek to understand all of these elements of the "intelligibility cocktail" (Zielinski, 2006) in part to inform pronunciation instruction and syllabus design, and our data do not yield enough consistent information to make unequivocal, or new, suggestions for prioritizing certain features of pronunciation. While the challenges of /æ/ merit further exploration, we were unable to find any patterns or problems that would indicate the need for pronunciation instructors to focus on specific vowels, consonants, stress patterns, etc.

Our findings on repair strategies may have more pedagogical implications. As noted above, NNSs made extensive use of repetition when they were misunderstood. It may be helpful to teach a variety of repair strategies in order to make learners aware of other options. A sample assignment can be found in Smith et al. (1992, pp. 23-26), in which learners practice using a variety of repair strategies through dialogues. Along the same lines, it may be beneficial for learners to be coached on when it is appropriate to abandon repetition as a repair strategy and move on to something else.

Another salient theme in our data that may have pedagogical implications is the emotional element to miscommunications. Our subjects reported reactions such as "frantic," "embarrassed," "shocked," and "confused" in response to misunderstood words or phrases (such reactions were more prevalent in NS – NNS exchanges, while NS – NS exchanges were more often laughed off). We saw self-reproach, especially among ESL teachers, at failures to understand: "I would have expected my ear to 'auto-correct' 'meth' to 'math' in a conversation where I was not yet listening specifically for pronunciation issues."

And there were emotions of triumph at resolutions: “At last the clerk understood what I wanted and I bought the watch. It did make my day.”

We therefore propose activities and strategies that would help alleviate some of the self-consciousness and stress involved in miscommunications. For example, for NNSs, pronunciation instructors may wish to share their own unintelligibility tales, in order to remind students that even proficient speakers have such experiences, and that they are not unique in their challenges. In addition, for NSs, Derwing, Rossiter, and Munro (2002) found that NS social work students who learned about typical English pronunciation difficulties of their Vietnamese clients often felt more secure in their interactions. Native speakers may increase their self-confidence by becoming aware of specific sounds that are challenging to NNSs with whom they interact.

CONCLUSION

This study enabled us to observe some of the drawbacks and benefits of using stories to elucidate the nature of (un)intelligibility.

There are various problems with using stories to examine exchanges involving unintelligibility. The biggest drawback is not having access to a recording of the exchange. For example, a recording of “Look at his tattoo” would enable us to examine the actual sounds produced and compare them with what the listener thought she heard. In addition, the stories we collected sometimes lacked information that would have been helpful: details about the repair, speaker or hearer proficiency, or the native language. Perhaps some of these gaps could be closed in follow-up interviews. Finally, our storytellers described misunderstandings rather than non-understandings. While our prompt clearly led them in that direction, we also realize that there are many cases “in real life” when exchanges result in no meaning being conveyed, as opposed to the wrong meaning. These are undoubtedly harder to capture, and harder to remember. And of course, they often don’t make a good story.

Beyond the challenges of the storytelling method, we have other limitations to our data. Foremost, a large proportion of our stories were from people with expertise in English language teaching, often with backgrounds in phonology. Thirty-four of our stories were from someone with expertise (most often the listener in the story). While these subjects’ stories were rich in detail and analysis, they are undoubtedly not representative of the kinds of stories that the general population would tell. In addition, we received very few stories from NNSs, and we would be keen to gather more in order to look for patterns and trends.

We also discovered many positive aspects to our methodology. The use of narratives or stories for research purposes has its roots in educational research (e.g. Clandinin & Connelly, 1989). The premise is that narratives “name a fundamental quality of experience, both personal and social” (Clandinin & Connelly, 1989, p. 4). Of interest is not so much what exactly happened, but how the storyteller “re-stories,” interprets, and makes meaning out of the experience. In this regard, stories of unintelligibility help us understand how listeners or speakers might bring their backgrounds, knowledge, and personal and social awareness to bear in an interpretation of a miscommunication event. In particular, we have seen how emotions play a role in miscommunications; this information may not be as easily uncovered in other research methods. Another revealing element of these stories was the descriptions of how the storytellers

sorted out what was wrong, or wrestled with how to interpret or deal with the situation – a "Here's what I was thinking" perspective. For example:

> At the beginning of class the next day, I asked the students what they thought about the first chapter of the book. One of my students, Julio (from Mexico), immediately raised his hand and said, "But Mrs. K., isn't there any alcohol in this book?" *I was confused. (I think that there is a character in the book who does some drinking, but I had no idea where this student was coming from with this question.)*
>
> "All day yesterday, you were talking about Tequila Mockingbird, and now I see it's To Kill a Mockingbird."
>
> At that point, the whole class laughed, and we moved into our discussion. However, *I remember thinking at the time about how funny his absolute disappointment was in discovering that there wasn't any alcohol in this whole book.* I hadn't had to do anything to repair the miscommunication myself—*I think Julio had thrown his book in his backpack at the end of class and pulled it out in his room that night to read the assigned chapter and probably figured out his misunderstanding on his own.*

Another advantage of collecting narratives is that it is possible to get stories from a multiplicity of contexts from around the world. The stories we received took place on nearly every continent, and in a range of social situations. Other data collection methods may not allow for such variety.

Finally, while we are uncertain about the usefulness of storytelling as a significant research method for understanding intelligibility, we are convinced of its worth as a pedagogical tool. In a pronunciation classroom, having students – and teachers – share their intelligibility tales can result in practical insights, analytical skill development, and perhaps lowered anxiety about "getting stuck" in such exchanges. At the same time, we would welcome further discussion among researchers and practitioners about this methodology and possibilities for refining it to capture some aspects of the "intelligibility cocktail."

NOTE

The use of more than one strategy was reported in a number of cases, leading to a total percentage over 100.

ABOUT THE AUTHORS

Laura Hahn is the Director of the Intensive English Institute at the University of Illinois at Urbana-Champaign. Her professional experience and research interests center on English pronunciation for international teaching assistants, ESL pronunciation pedagogy, and teacher development. She is the co-author, with Wayne B. Dickerson, of *Speechcraft: Discourse Pronunciation for Advanced Learners* (1999, University of Michigan Press).

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.

Contact information:
Laura Hahn
University of Illinois at Urbana-Champaign
Intensive English Institute, 1205 W. Nevada St., Urbana IL 61801
217-333-6598

Patricia Watts
University of Illinois at Urbana-Champaign
Department of Linguistics, 707 S. Mathews Ave., Urbana IL 61801

217-333-1506 pawatts1@Illinois.edu

REFERENCES

Bond, Z. (1999). *Slips of the ear: Errors in the perception of casual conversation.* San Diego: Academic Press.

Clandinin, J. & Connelly, M. (1989). Narrative and story in practice and research. ERIC Document Report 309681.

De Decker, P., and Nycz, J. (2005). What lies beneath? Studying variation using ultrasound technology. Paper presented at the 12th International Conference on Methods in Dialectology, Université de Moncton, Canada.

Derwing, T., Rossiter, M., & Munro, M. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingual and Multicultural Development 23*(4), 245-259.

Derwing, T. & Rossiter, M. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System* 30(2), 155-166

Smith, J. Meyers, C, & Burkhalter, A. (1992). *Communicate: Strategies for international teaching assistants.* Waveland Press, Long Grove, IL

Zielinski, B. (2006). The intelligibility cocktail: An interaction between speaker and listener ingredients. *Prospect 21*(1), 22

APPENDIX A: WRITING PROMPT

What's Your Unintelligibility Tale?

Here is a condensed example of an unintelligibility tale:

A US American female is with her friends at a Korean restaurant in the United States. They have ordered several dishes to try. The Korean waiter brings out the dishes.

Waiter: Do you like pork?

Customer: (trying to remember what they ordered) Yes.

The waiter disappears and comes back with forks for all of the people at the table.

These exchanges and our reactions to them tell us a great deal about how intelligibility works.

Directions
Write a **narrative** describing a miscommunication that took place due to a pronunciation misunderstanding you experienced, with either a native speaker or a non-native speaker of English.

Your narrative should include four to five short paragraphs that include the following information:

- Participants
  Who participated in this conversation?
  What are their native languages (if you know)?

  How are the participants related to each other (friends, teacher-student, etc.)?
  Male or female?
  Are there any other relevant details about the participants?
  Context

  Where and when did the exchange take place?
  What was the purpose of the exchange (e.g. small talk, asking for help)?

- Dialog
  Write down the dialog as best as you can remember it.

- Comments and interpretation
  What you were thinking as you experienced the exchange?
  How did you become aware of the confusion?
  What were you thinking about how to repair the miscommunication?

- Other information
  Please include any other information that will help us understand what happened.

APPENDIX B: MISUNDERSTANDINGS

Case (intended meaning, interpretation)	**Who misunderstands whom?**
Mac, Matt	NNS - NNS
tofu, TOEFL	NNS - NNS
grandma, grammar	NNS - NNS
marmots mormons mammals	NNS - NS
condo condom	NNS - NS
small talk, sumo talk	NNS - NS
To Kill a Mockingbird, tequila mockingbird	NNS - NS
program, problem	NNS - NS
car heart	NNS - NS
tutor, torture	NS - NNS
tattoo, statue	NS - NNS
duck, dog	NS - NNS
dog eat dog, doggie dog	NS - NNS
McDonald's, unintelligible	NS - NNS
math camp, meth camp	NS - NNS
ranking, linking	NS - NNS
register, REEject	NS - NNS
one inch, one in	NS - NNS
fork, pork	NS - NNS
insurance, insulin	NS - NNS
user name, your surname	NS - NNS
Penfold wines, Ben Folds Five	NS - NNS

snap, snuff	NS - NNS
important, impotent	NS - NNS
city, seetch	NS - NNS
Busey, bush	NS - NNS
Baytown, Daytown	NS - NNS
passion, patience	NS - NNS
bed time, bad time	NS - NNS
twin bedroom, two in bad room	NS - NNS
ring/rain	NS - NNS
What does Atsuko do?, When is Atsuko due?	NS - NNS
you'll, you	NS - NS
meat, beef	NS - NS
soup or salad, super salad	NS - NS
taxes, Texas	NS - NS
Al and Prasanta, Alan Prasanta	NS - NS
gorgeous sky, gorgeous guy	NS - NS
ínvalid diet, inválid diet	NS - NS
She's in Ames/She's engaged	NS - NS
in-house test, enhanced test	NS - NS
soup or salad, super salad	NS - NS

Algethami, G. Ingram, J., & Nguyen, T. (2011). The interlanguage speech intelligibility benefit: The case of Arabic-accented English. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 30-42), Ames, IA: Iowa State University.

THE INTERLANGUAGE SPEECH INTELLIGIBILITY BENEFIT: THE CASE OF ARABIC-ACCENTED ENGLISH

Ghazi Algethami,[1] University of York

Dr. John Ingram, University of Queensland

Dr. Thu Nguyen

This study attempted to further examine the so-called *Interlanguage Speech Intelligibility Benefit*, which refers to the intelligibility advantage L2 listeners have over native listeners when they listen to L2 speech produced by speakers who share the same native language. 19 native speakers of Australian English and 19 L2 speakers of English whose native language is Saudi Arabic listened to English utterances produced by ten L2 Saudi speakers of English who were of two groups; high pronunciation proficiency and low pronunciation proficiency. Utterances from three native speakers of Australian English were also included as controls. The listeners were asked to transcribe what they thought the speaker had said in English orthography. The percentage of words identified correctly by each listener for each utterance is taken as the intelligibility score. Although the L2 Saudi listeners did slightly better than the native listeners when they listened to Saudi accented English, both listener groups did not differ significantly. The results corroborate previous findings suggesting that the intelligibility advantage for the non-native listeners is small, if any, and the phonetic properties of the L2 speech itself are strong determinants of how it is perceived regardless of the listeners' native language.

INTRODUCTION

It has long been suggested that non-native listeners have an intelligibility advantage over native listeners when they listen to non-native speech produced by L2 speakers who share the same L1 (Weinreich, 1953). This means that the non-native listeners could do better than the native listeners at understanding L2 speech produced by those who share with them the same L1. Recent studies have attempted to test this intuition empirically (Bent & Bradlow, 2003; Hayes-Harb, Smith, Bent, & Bradlow, 2008; Imai, Walley, & Flege, 2005; Ingram & Nguyen, 2007; Major, Fitzmaurice, Bunta, & Balasubramanian, 2002; Munro, Derwing, & Morton, 2006; Stibbard & Lee, 2006; van Wijngaarden, 2001; van Wijngaarden, Steeneken, & Houtgast, 2002). Generally, these studies found very small and inconsistent evidence for this advantage.

For example, Stibbard and Lee (2006); Major et al. (2002); and Bent and Bradlow (2003) showed that L2 listeners did not outperform native listeners in recognizing words produced by L2 speakers who share the same L1 background. In contrast, Imai et al. (2006) reported that L2 Spanish listeners yielded higher intelligibility scores than native listeners when they were both presented with Spanish-accented English. Munro et al. (2006) found that while the Japanese listeners found Japanese-accented English more intelligible than did the native listeners, the

[1] Authors ordered by percentage of contribution to the current study.

Cantonese listeners did not report such an advantage when listening to Cantonese accented English.

It was also suggested that this intelligibility advantage might be mediated by the pronunciation proficiency level of the L2 speakers, which has been mostly operationalized as the degree of foreign accent they speak with (Hayes-Harb et al., 2008). For example, Hayes-Harb et al. (2008) showed that low proficiency L2 Mandarin listeners did better than high proficiency L2 Mandarin and native English listeners when they listened to speech produced by low proficiency Mandarin speakers.

Bent and Bradlow (2003) called this advantage *The Interlanguage Speech Intelligibility Benefit*. However, it should be noted that they used it to mean that the native and non-native listeners were equal in understanding the non-native speech. Stibbard and Lee (2006) argued, however, that the non-native listeners have to outperform their native counterparts in understanding the L2 speech to be called an advantage. The current study follows Stibbard and Lee.

A major limitation in most previous studies is that they used speech which was elicited from L2 speakers by explicit tasks of reading and pronunciation assessments, a method which would encourage the L2 speakers to monitor their production and hence undermine some of the L1 interference, which has been assumed to enhance the intelligibility for the non-native listeners who share the same L1. However, reliance on read speech is almost inescapable as it offers control over the content to be compared. An exception to this limitation is Munro et al. (2006). They extracted utterances from narratives obtained by asking participants to describe extemporaneously a cartoon story. Nevertheless, this method does not offer us complete control over what lexical items would be included in the utterances to be presented to the listeners. Thus, differences between intelligibility scores given to the listeners might be affected by the differences between the utterances. In addition, Munro et al. (2006) did not examine how the proficiency level of the L2 speakers or the degree of their foreign accents affects their intelligibility to the L2 listeners who share the same L1.

The current study implemented an elicitation method, first used in Ingram and Thu (2007), to both offer control over the content or the lexical items to be included in the utterances and to deflect the L2 speakers from monitoring their L2 production. The speakers were asked to paraphrase a number of English sentences within a specified time-duration, and then write and read out their responses into a microphone. The purpose of the time-allotted paraphrase task was to place a moderate cognitive load on the L2 speakers so that they would be preoccupied with formulating the sentences rather than with monitoring their pronunciation. It also offered a way to control the lexical items to be included in the listening task. The complexity of the paraphrase task was sufficient to engage the L2 speakers and deflect their attention from the pronunciation side of the task.

Ingram and Thu (2007) did not find an intelligibility benefit for L2 Vietnamese learners of English when they listened to Vietnamese accented English, but rather a disadvantage as the native English listeners found the Vietnamese accented English more intelligible than did the Vietnamese L2 listeners. However, they did not examine the effect of the proficiency level of the L2 speakers which has been shown to affect the degree of intelligibility among listeners (Hayes-Harb et al., 2008).

The current study examines whether L2 Saudi listeners outperform native English listeners in understanding Saudi Arabic-accented English. A few studies have examined the production of English speech by L2 Saudi speakers. Flege and Port (1981) studied the production of English stop consonants by L2 Saudi speakers. Their results generally showed that the L2 Saudi speakers used their L1 Arabic temporal acoustic correlates of stop consonants when they were asked to produce English stop consonants. However, except for /p/, they showed that this acoustic deviation from the native norms did not pose much difficulty for native listeners when they were asked to identify which stop consonant was produced. Munro (1993) studied the production of English vowels by 23 L2 Arabic speakers, of whom only three were Saudis. He pointed out that "the acoustic properties of Arabic-accented vowels were: exaggerated tense-lax duration differences, small voicing-conditioned duration differences, a very short /ɑ/, a tendency toward a low F1 in /ɛ/ and /ɑ/, a tendency toward a low F2 in most back vowels, and generally less formant movement in F1 and F2 than in the native speakers' productions" (p. 51). However, it is somewhat difficult to attribute these acoustic properties to the Saudi-accented English because the participants were of different Arabic dialects which are phonologically different (Watson, 2002). Other studies have used only auditory analysis to examine the English pronunciation mistakes made by L2 Saudi speakers. Altaha (1995) pointed out some phonemes which were difficult for his students to produce: /p/, /v/, /ɛ/, /ɜ:/ and /æ/. Basalamah (1990) showed that his L2 Saudi participants used their L1 Saudi Arabic stress pattern when speaking English. Except for Flege and Port (1981), none of these studies has shown how these deviations from the native norms observed in the speech of L2 Saudi speakers would affect their intelligibility. It is also worth pointing out that deviations from the native phonetic norms perceived in the speech of second language speakers could cue our perception of foreign accentedness, but they might not necessarily reduce its intelligibility (Derwing & Munro, 1997). Therefore, it is still not clear what pronunciation mistakes or deviations from the native speech would make the L2 speaker harder to understand. The investigation of these mistakes is beyond the scope of the current study, though.

Before describing the study, some related terminological issues should be explained. Derwing and Munro (1997) defined *intelligibility* as the actual understanding of the message. Most previous studies measured *intelligibility* by counting the number of correct words identified by the listeners. However, the correct identification of words does not necessarily mean *understanding the message* as listeners might be able to identify all the words while still puzzling over what the speaker is trying to communicate (Zielinski, 2004). For convenience and consistency with previous studies, the current study uses *intelligibility to* mean the correct identification of the words uttered by the speaker.

The pronunciation proficiency level of L2 speakers was used in some previous studies to mean the degree of foreign accent they speak with (Hayes-Harb et al., 2008; Imai et al., 2005; Stibbard & Lee, 2006). In these studies, L2 speakers were grouped into two subgroups according to the degree of foreign accent. Generally, those with stronger foreign accents were categorized as *Low Proficiency* L2 speakers and those with less foreign accent were categorized as *High Proficiency* L2 speakers. However, the term *proficiency* in pronunciation might also entails other dimensions that are not directly related to foreign accentedness, such as fluency and grammaticality (Derwing & Munro, 1997). For the sake of consistency with previous studies, the current study uses the terms *high proficiency* and *low proficiency* to refer to the degree of foreign accentedness, as measured by the ratings given by the native listeners in this study.

METHOD

First, a description of how the stimuli used in the listening test were constructed is given. Then, the construction and delivery of the listening test which provided the data for the current study are described.

Stimulus Construction

Twenty three English sentences were adopted from Ingram and Thu (2007) (see Appendix A). There was no basis for choosing these utterances other than that of convenience.

Speakers

The speakers who supplied the listening materials for the present study were full-time international Saudi postgraduate students at Australian universities in Brisbane. They were in the age range of 20-35 and ranged in residency in Australia from a period of 1 to 4 years. All of them had attained written and spoken English proficiency that enabled them to undertake Master and PhD degrees at the University of Queensland, Griffith University and Queensland University of Technology. Three native speakers of Australian English were also recruited to provide native English utterances to be used as controls.

Speech Elicitation Method

The current study attempted to create a speaking situation where there is a moderate cognitive load involved and where the speakers are more preoccupied with the formulation of the message rather than with self-monitoring of their pronunciation. A grammatical paraphrase task was found to meet these requirements and at the same time provide control over the lexical selection of items among the speakers.

Grammatical Paraphrase Task

The grammatical paraphrase task required subjects to transform sentences, which were constructed previously, presented in spoken and written form (over headphones and a computer screen) into a meaning-equivalent form. The materials were presented via a spoken Language Assessment Program (Harrington & Ingram, 2003). The participants typed in the paraphrase in response to an initial prompt word and when satisfied with their construction, read out the sentence that they had formed. The linguistic aspects of the task were sufficiently complex to engage the subjects (who were all L2 speakers of English, with the exception of three controls) and to deflect their attention from the pronunciation aspect of the task. A sample of this production is presented below:

Example Stimulus: *The soldier's face and mouth were covered by a mask.*

(Audio + written)

Visual Prompt: *A mask ……*

Paraphrase response: *A mask covered the soldier's face and mouth (*typed, and then spoken).

After speaking their paraphrase responses into a headset microphone, subjects pressed a button for presentation of the next item in the set, randomly selected without replacement until all the items had been presented. The typed response and the audio signal were saved to a database from which the listening items were selected. Because some of the L2 speakers were not able to come to the phonetic lab at the University of Queensland, and access to the Language Assessment Program (Harrington & Ingram, 2003) was not possible, recording took place in a quiet furnished room in their homes using a personal laptop. The paraphrase test was presented via a word document, where the participants were asked to click on each sentence to hear it, paraphrase it according to a prompt word, and then read it aloud into a microphone. The participants were told that they had thirty minutes to complete the test.

Construction of the Listening Test

Most of the participants for the listening test were drawn from an introductory linguistics class at the University of Queensland. They participated in the listening experiment for course credit. More L2 Saudi participants in the listening test were needed. They were drawn from the Saudi Students Society in Brisbane. All of them were full-time postgraduate students at the University of Queensland, Griffith University and Queensland University of Technology. Their range of residency in Australia was from 1 to 4 years. The experimental task was time-constrained (approx. 20 minutes) to avoid fatigue and flagging interest.

Due to likely practice effects if any test utterance was heard repeatedly in the course of the experiment, it was necessary to block items in such a way that no listener heard a given utterance more than once (with the exception of four control utterance at the end). It was also necessary to expose each listener to the full range of speaker variation in the test sentence set.

For this reason, each listener was assigned to one of 18 overlapping blocks of 23 items. The four control items (spoken by native speakers of Australian English) which involved repetition of test utterances were always presented as the last four items in a block, so the perception of accented English items would not be affected by previous exposure to the same utterance produced by different speakers.

Subjects for the Listening Experiment

The first group of listeners was comprised of native-born speakers of Australian English (N= 19), who reported no familiarity with Arabic-accented English. The second group of listeners was made up of Saudi students whose native language is Saudi Arabic (N=19).

Procedure for the listening experiment

Participants in the listening experiment were randomly and evenly allocated to one of 18 blocks of items (versions of the experiment). As there were 19 listeners of each group and 18 blocks of utterances, two listeners in each group had to listen to the same block of utterances. The eighteen overlapping blocks of the listening test were needed because no listener could hear a given utterance more than once in an attempt to expose each listener to the full range of speaker variation in the test sentence set and to avoid the effect of repetition in the listening experiment. A version of the experiment and instructions for undertaking it were e-mailed to the subjects as word-files containing links to sound-files for each utterance uploaded in the University of

Queensland web-domain (Appendix B). Subjects had the option of doing the experiment at their time of choosing, using the multi-media facilities of their home computers or using a University machine. Response sheets were e-mailed to the experimenters.

Subjects were instructed that they could play each item up to four times. They were told to transcribe what they thought the speaker said into standard orthography and then to rate the original utterance for accentedness on a 5-point scale: (1) no detectable foreign accent, (2) mild accent, (3) moderate accent, (4) strong accent, (5) very strong accent.

RESULTS

The percentage of words identified correctly by each listener for each utterance was taken as the intelligibility score. Errors including word omissions, word substitutions, and incorrectly transcribed inflections, but excluding spelling errors, were also counted but not included in the analysis for this study. The L2 Saudi speaker group was divided into two groups- high proficiency group and low proficiency group- based on mean ratings of the degree of foreign accent given to them by the native listeners. L2 speakers who received mean rating of 2.5 or less were considered high-proficient while those who received mean rating above 2.5 were considered low-proficient. The non-native listeners' ratings of foreign accentedness were not taken into account when dividing the L2 speakers into high and low proficiency groups as the reliability of their ratings might be affected by their L1. Figure 1 shows the mean foreign accent ratings given to both the native and L2 speakers by the native listeners.

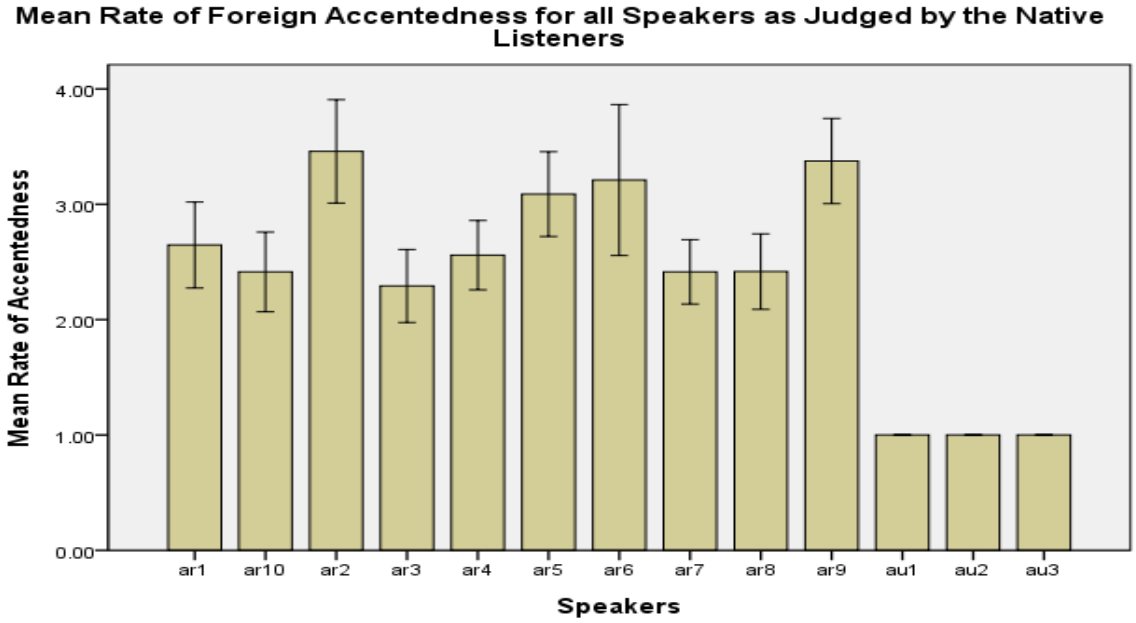


Figure 1. Mean rating of the degree of foreign accent for both the native speakers (au) and the L2 speakers (ar).

Unless stated, the significance level reported here is $p < .05$. A two-way analysis of variance with listener group (two-levels: Saudi and English) as a between-subject factor and speaker

group (Native: N; High Proficient: HP; and Low Porficient: LP) as a within subject factor showed a significant main effect of speaker group ($F(2,873) = 7.574$, $p < .05$), and no significant effect of listener group ($F(1,873) = 2.846$, $p > .05$). However, there was a significant interaction between listener group and speaker group ($F(2,873) = 5.682$, $p < .05$).

Post-hoc pairwise t-tests showed no significant difference between the two listener groups when they listened to both high and low proficiency L2 Saudi speakers of English. In contrast, the two listener groups differed significantly when they listened to the native English utterances. Table 1 shows the average intelligibility scores given for each listener group when listening to both the native and non-native speech. Figure 2 also shows the results graphically.

Table 1. *Average intelligibility scores given for each listener group*

Listener	Speaker	% Intelligibility Score
Saudi	LP	91.722
	HP	94.494
	Native	92.533
Native	LP	90.478
	HP	93.752
	Native	100.000

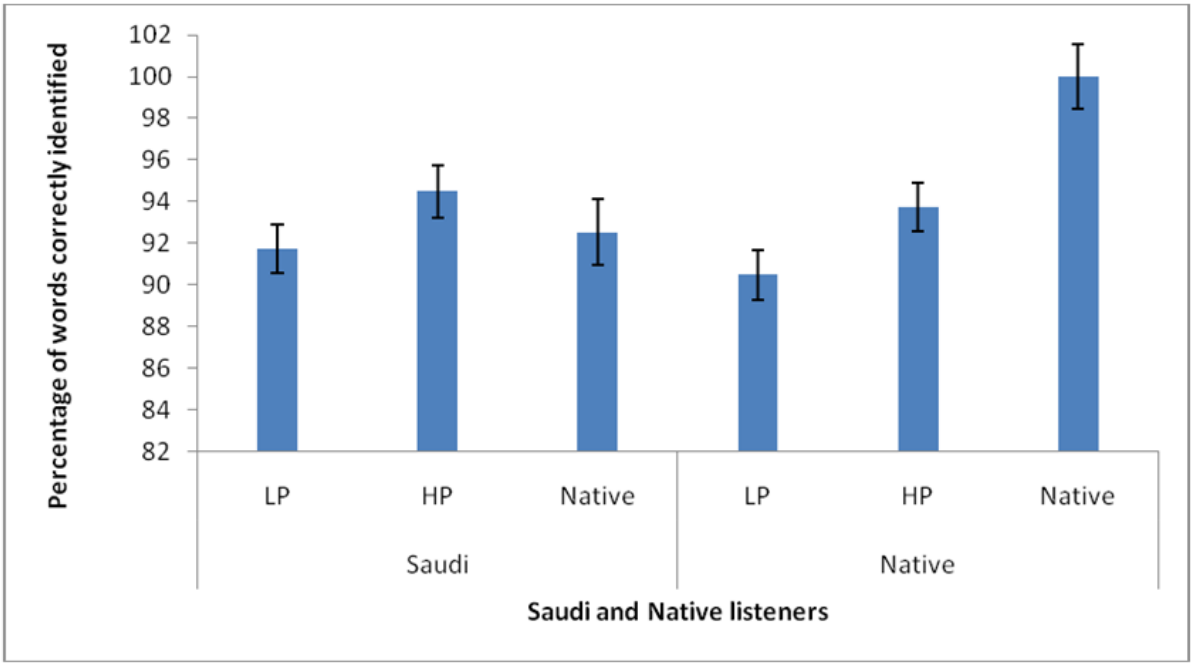


Figure 2. Average intelligibly score for both listener groups for each speaker group.

DISCUSSION AND CONCLUSION

Previous studies found small or no intelligibility advantage for non-native listeners over native listeners when they listened to L2 speech from speakers who share the same L1. It has been suggested that the L2 speakers who share the same native language also share the same interlanguage, a fact assumed to offer them an intelligibility advantage over the native listeners (Bent & Bradlow, 2003). We hypothesized that previous studies did not find such an advantage because the speech stimuli used in the listening experiments were elicited by explicit ways of reading and pronunciation assessments. The current study used a methodology to try to deflect the L2 speakers from monitoring their speech, a fact that might undermine some of the L1 interferences which might facilitate the intelligibility of L2 speech for L2 listeners who share the same L1.

Although the non-native listeners yielded higher intelligibility scores than the native listeners for both L2 speaker groups, the difference was quite small and did not reach statistical significance. This gives further credence to the previous studies that showed small or no intelligibility benefit for non-native listeners when they listened to L2 speech produced by L2 speakers who share the same native language.

One might conclude that despite the degree of foreign accent L2 speakers have, there are certain aspects of L2 speech that seem to be detrimental to intelligibility for both native and non-native listeners. Future research on L2 speech must seek what these features are (Munro, 2008). These features will be appreciated by L2 teachers and curriculum designers as intelligibility places a high priority in teaching and learning a second language.

Although it is not the purpose of the current study, it might be useful to mention some of the pronunciation mistakes made by the L2 Saudi speakers which impeded intelligibility. We noticed that they, especially the low proficiency group, tended to produce /ɒ/ as /u/ or sometimes /o:/. Both L2 groups showed less diphthongization than the native speakers when they produced diphthongal segments. The stop consonant /p/ was in many cases produced as /b/. Stress placement and timing also played a role in reducing the intelligibility of the L2 speech.

ABOUT THE AUTHORS

Ghazi Algethami is a PhD candidate at the University of York. His general research interest is the production and perception of second language speech.

Affiliation: University of York.
Address: United Kingdom, York, Heslington, YO10 5DD.
Email: ghazi.algethami@gmail.com

Dr. John Ingram is a senior lecturer at the University of Queensland. His research interests include Phonetics, Phonology, Second language speech, Psycholinguistics, and Neurolinguistics.

Affiliation: University of Queensland.
Address: Australia, University of Queensland, School of Languages and Comparative Cultural Studies, St Lucia, 4072.
Email: j.ingram@uq.edu.au

Dr. Thu Nguyen's research interests include second language speech, Vietnamese accented English, L2 prosody, and Vietnamese phonology.

Email: nguyenthianhthu@email.com

REFERENCES

Altaha, F. (1995). Pronunciation errors made by Saudi university students learning English: analysis and remedy. *International Review of Applied Linguistics, 109-110*, 110-123.

Basalamah, M. (1990). *A Study of Certain Aspects of the Suprasegmentals of Arabic and their Influence on the English Pronunciation of Arab Teachers of English in the Western Zone of Saudi Arabia*. Unpublished PhD thesis, University College London, London.

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America, 114*, 1600-1610.

Derwing, T. M. & Munro, M. J. (1997). Accent, comprehensibility and intelligibility: Evidence from four L1s. *Studies in Second Language Acquisition, 20*, 1-16.

Flege, J. E. & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech, 24*(2), 125-146.

Hayes-Harb, R., Smith, B., Bent, T., & Bradlow, A. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin:Production and perception of English word-final voicing contrasts. *Journal of Phonetics, 36*, 664-679.

Imai, S., Walley, A. C., & Flege, J. E. (2005). Lexical frequency and neighbourhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Journal of the Acoustical Society of America, 117*, 896-907.

Ingram, J. & Nguyen, T. (2007). *Intelligibility of Vietnamese accented English.* Paper presented at the TESOL in the internationalization of higher education in Vietnam, Hanoi, Vietnam.

Harrington M. & Ingram J. (2003). LanguageMAPTM: A multi-media language assessment program, http://www.languagemap.com.

Major, R., Fitzmaurice, S., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment *TESOL Quarterly, 36*(2), 173-190.

Munro, M. (2008). Foreign accent and speech intelligibility. In J. Edwards & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 193-218). Philadelphia: John Benjamins.

Munro, M., Derwing, T., & Morton, S. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28,* 11-131.

Stibbard, R. M., & Lee, J.-I. (2006). Evidence against the mismatched interlanguage intelligibility benefit hypothesis. *Journal of the Acoustical Society of America, 120,* 433-442.

van Wijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication, 35,* 103-113.

van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of the Acoustical Society of America, 111,* 1906–1916.

Watson, J. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Zielinski, B. (2004). *Measurement of speech intelligibility: What are we actually measuring?* Paper presented at the Conference of American Association for Applied Linguistics, Portland, USA.

Appendix A

Australia is the world's driest continent
It is cold soup
Jack tried to steal the jewels
Ann will paint the room
The strong box was where john kept his money
The window will be washed by Nick
Because it was dark John stumbled at the blackboard
Elephants used to be hunted for their tusks and hides
John was admired by the other students
The room will be painted by Ann
Without his glasses, John could not see the blackboard
Alan folded the newspapers
Kim wanted a box of chocolate, not a bunch of flowers
The other students admired John
Dark clouds were gathering as they came across the bridge
He needs a strong box to carry all those books
The bluebells and tulips were hung from the eaves of the green house
The soldiers face and mouth were covered by a mask
The plate ran away with the spoon
Bill's life savings were invested in a friendly society
Where the queen was sleeping was the royal tent
What they wanted was to migrate to a friendly society

Appendix B

Accented English Experiment - Group 1 (version 1)

The aim of this experiment is to assess the effects of non-native accent upon spoken sentence comprehension. We are interested in finding out how the strength of a non-native accent is related to how easy/difficult particular spoken sentences are to understand. Please listen to each sentence (up to, but **not** more than **four** times). Type what you think that the speaker said in the space provided. Then rate the sentence for the strength of the speaker's non-native accent. Type in the words you think you recognize even if the sentence as a whole may not make sense to you.

Your last name:	**First name**:
Student number:	**Gender (M, F)**: **Age:**
Your native language:	**Mother's native lang.:**
Father's native lang.:	**Years lived in Australia:**
Other languages you speak/understand well:	
Places you have lived for more than 6 months outside Australia:	
Main language spoken at home:	
Accented Englishes that you are familiar with hearing (check with X): Chinese ( ) Thai ( ) Vietnamese ( ) Japanese ( ) Indian ( ) Arabic ( ) Other ( ) None of these ( )	

Please fill in the following information:

Instructions:

The first column of each row in the table below contains an **Item,** a short spoken sentence. If you hold down the **Ctrl** key (bottom left of your keyboard) and move the mouse cursor over the item a small hand will appear. Right click on the cursor to hear the item. You may play each item *up to four* (**4**) times, but **not** more than that. When you have finished listening to an item, close the 'sound play' window and answer the following questions on the questionnaire:

In the first column type in what you think the speaker said, or at least those words that you thought you heard, even if the sentence as a whole does not make sense.

In the second column make a rating as to how strong the speaker's accent appeared to be:

Strength of non-native English accent

(1) no foreign accent (2) mild accent (3)moderate accent (4) strong accent (5) very strong accent

When doing this experiment, please work entirely on your own. Work through each item in the order given in the table. Having completed an item, move to the next one. Do not go back and change any of your answers. Do not play any item more than four times. It is your first listening impression that we are after. So do not spend much time on any given item.

Item No.	**Type what you think that the speaker may have said**.	**Accent Rating**
item01		
item02		
item03		
item04		
item05		
item06		
item07		
item08		
item09		
item10		
item11		
item12		
item13		
item14		

Item No.	**Type what you think that the speaker may have said**.	**Accent Rating**
item15		
item16		
item17		
item18		
item19		
item20		
item21		
item22		
item23		

Thank you!

When you have finished filling in this form, please save it with a file name that includes your name and student id: (e.g.: 'accent_xpt_Joe_Bloggs_129457').

Having saved the file, please send it back as an email attachment.

Lima, E. F. (2011). Language and nonlanguage factors affecting nonnative undergraduate students' reactions to ITAs. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (pp. 43-55), Ames, IA: Iowa State University.

LANGUAGE AND NONLANGUAGE FACTORS AFFECTING NONNATIVE UNDERGRADUATE STUDENTS' REACTION TO ITAS

Edna F. Lima, Iowa State University

This study investigates whether language and nonlanguage factors affect international undergraduates' perceptions of international teaching assistants (ITAs). Fifty-five students enrolled in first-year composition classes watched a short video-taped lecture under one of three guises related to nationality of the speaker and rated the lecture and the speaker based on eight response variables. Results indicate that the information provided to participants about the nationality of the speaker did not influence their perception of both lecture and speaker. However, when participants' variables were analyzed, statistically significant results were found for two response variables: accent and speaker likeability. The results for accent indicate that the actual degree of accentedness that participants perceived in the speaker's speech, not nationality, influenced their ratings. As for likeability of the speaker, raters favored the supposed Brazilian TA. This finding may be related to stereotypes of Brazilian people and culture worldwide or to previous socio-cultural experiences that participants may have had with Brazilian individuals.

INTRODUCTION

As the number of international teaching assistants (ITAs) in the U.S has increased throughout the years, there has been a growing concern about communication between undergraduate students and their ITAs both in classrooms and in office hours (Damron, 2003). Davis (1991) argues that lack of oral proficiency and cultural differences are generally judged to be two major problems that ITAs face when teaching at American universities. He asserts that ITAs cannot communicate effectively with students in the classroom due to their limited oral proficiency. Davis also claims that the interaction between instructors and students is not effective because of different expectations regarding the role of both instructors and students and the goals and processes of higher education.

Native undergraduate students' reaction to ITAs

Research has shown that several are the factors influencing American undergraduates' perceptions of ITAs. Most complaints from native undergraduates about ITAs concern poor English language proficiency and/or communicative competence (Lindemann, 2002; Orth, 1982; Rubin & Smith, 1990; Rubin, Ainsworth, Cho, Turk, & Winn, 1999); however, studies have revealed that factors other than linguistic reality play a bigger role in this perception. Among such factors we may find age, gender, country of origin, teaching style, cultural background, personality, first language, accent, topical knowledge, and so on (Brown, 1992; Gill, 1994; Rubin, 1992).

Accentedness is often regarded by native speakers as the biggest constraint in their communication with nonnative speakers; nonetheless, research findings have shown that listeners will rate some speakers' utterances as heavily accented even though these utterances are perfectly intelligible and totally comprehensible (Munro & Derwing, 1995). For instance, Rubin (1992) investigated 62 North American students' perception of instructors using a picture guise; a group of participants listened to a lecture in conjunction with the picture of an Asian TA; another group listened to the lecture in conjunction with the picture of a Caucasian/European TA; and another group listened to the lecture without any photographs. The results indicated that lower teacher effectiveness ratings were assigned for speakers receiving more negative accent ratings. The accent was perceived as more foreign and less standard for the Asian instructor's photograph. Therefore, perceived accentedness, not actual accent, was negatively connected to instructor ratings. Impatience, inexperience with L2 speakers, and prejudice are some of the factors leading listeners to react negatively to accented speech (Lippi-Green, 1997; Munro, 2003; Munro, Derwing, & Morton, 2006).

Nonnative speakers' perception of nonnative speech

If native speakers tend to react negatively to foreign-accented speech, how do nonnative speakers react to the speech of other nonnative speakers? Munro, Derwing, and Morton (2006) argue that NNSs' responses to NNSs' utterances may vary depending on the degree of familiarity with or exposure to accents or on the listener's first language. Bent and Bradlow (2003) claim that nonnative listeners may regard foreign-accented speech more intelligible than native speech, and that the opposite may be true for native listeners.

Research on nonnative listeners' perception of native and nonnative speech is divergent. Some studies suggest that listeners from different L1 backgrounds show moderate to moderately high correlation in their responses. L1 background and experience with a given accent appear to be minor factors in the ability to understand L2 speech (Munro et al., 2006). Other studies suggest that for nonnative listeners, the intelligibility of high proficiency speakers from the same L1 background is similar to the intelligibility of native speakers and that the speech intelligibility of nonnative speakers from different L1 backgrounds is equal to or greater than the intelligibility of native speakers (Bent & Bradlow, 2003). On the other hand, research suggests that when it comes to nonnative students' preferences for native speaker teachers (NST) versus nonnative speaker teachers (NNST), 60.6% of the participants prefer NST, 35% show no preference, and only 3.9% prefer NNST. It is noteworthy, however, that when offered the option of a team-teaching approach (NST and NNST), 71.6% of the participants think it is a good idea (Lasagabaster & Sierra, 2005).

Given the lack of substantial research dedicated specifically to explore nonnative undergraduate students' perceptions of ITAs, this study investigates how nonnative undergraduate students react to ITAs with emphasis on both language and nonlanguage factors.

The 55 participants involved in the study were divided into three treatment groups and asked to watch and rate a lecture and its speaker. The eight response variables were *accent*, *speed*, *comprehensibility* (language factors), *level of interest in the lecture*, *usefulness of the lecture*, *likeability of the speaker*, *teaching ability of the speaker*, and *teaching style of the speaker* (nonlanguage factors). The predictor variables were *attributed nationality of the speaker*, *raters'*

gender, *raters' first language (L1)*, and *number of previous ITAs*. Although the lecture was exactly the same for all three groups, each group was given different information about the speaker. Group I was told that the speaker is an Egyptian teaching assistant, Group II received the information that the speaker is a Brazilian teaching assistant, and Group III was given the information that the speaker is an American teaching assistant. The research questions for the study are as follows:

1. Do the three groups rate the lecture differently depending on what they are told about the speaker? If so, in what specific areas do the groups rate the lecture differently?

2. How do the ratings differ across groups depending on raters' gender, first language (L1), and number of international teaching assistants they had class with before the study?

METHOD

Participants

Speaker

The speaker was a Serbian Ph.D. student enrolled in the Applied Linguistics and Technology (ALT) program at Iowa State University at the time the study was conducted. She was also a teaching and research assistant in the English Department. In addition to her near native-like pronunciation, this specific speaker was chosen because of her physical appearance, a key aspect in this study; it was crucial that the participants in each of the three treatment groups found the information about the speaker to be at least plausible.

Raters

The 55 raters taking part in this study were international students enrolled in two cross-cultural sections (roughly 50% of Americans and 50% of international students) of English 150, a writing foundation course for first-year undergraduate students, and in three cross-cultural sections of English 250, a writing course for second-year undergraduate students. The 55 participants were from 12 different countries: China (11), Ecuador (2), India (4), Indonesia (4), Japan (1), Korea (5), Libya (1), Malawi (1), Malaysia (20), Mexico (2), Saudi Arabia (1), and the United Arab Emirates (3). Thirty six of these participants were males and nineteen were females with ages ranging from 18 to 24, with 20.3 being the average age (SD 1.5). Their length of residence in the United States ranged from three months to nine years with an average of 10 months.

Given that the data collection process took place during regular class periods, it was impractical to deliberately assign participants to each of the three treatment groups. Thus, the number of participants was defined by the section of English 150 or 250 in which they were enrolled. Table 1 summarizes the demographic data of each group.

Table 1: Demographic data of the treatment groups

Treatment Group	N**	Gender	Age Average	Country of Origin	L1
*Group I (Egyptian TA)**	20	5 females, 15 males	20.4	China (2), Ecuador (2), Indonesia (2), Libya (1), Malawi (1), Malaysia (9), Mexico (1), Korea (1), Saudi Arabia (1)	Chinese (China), Spanish (Ecuador and Mexico), Indonesian (Indonesia), Arabic (Libya and Saudi Arabia), Chichewa (Malawi), Malay and Chinese (Malaysia), Korean (Korea)
*Group II (Brazilian TA)**	23	10 females, 13 males	19.8	China (9), India (4), Indonesia (2), Korea (2), Malaysia (2), UAE (2), Japan (1), Mexico (1)	Chinese, Hindi and Urdu, Indonesian, Korean, Malay and Chinese, Arabic, Japanese, Spanish
*Group III (American TA)**	12	4 females, 8 males	21.2	Korea (2), Malaysia (9), UAE (1)	Korean, Chinese and Malay, Arabic

* Information provided to each treatment group though the lecture and the speaker were exactly the same for all groups
** Number of participants per group

Materials

Personal Information Questionnaire

The first questionnaire designed for this study contained 11 questions and asked participants to provide their assigned ID number, age, gender, country of origin, native language, English proficiency level, educational background, foreign languages (other than English), major, length of residence in the United States, and the number of international assistants they had had class with prior to the study.

Video-taped Lecture

The stimulus was a four-minute video-taped lecture (Figure 1) recorded in a real classroom to keep authenticity of environment. Although the speaker was provided with a lecture script, it was essential that she applied her personal teaching style to the lecture and maintained her normal teaching speed. The topic chosen for the lecture was "thesis statement" because it is a topic pertaining to the courses that the participants were taking at the time of the study.

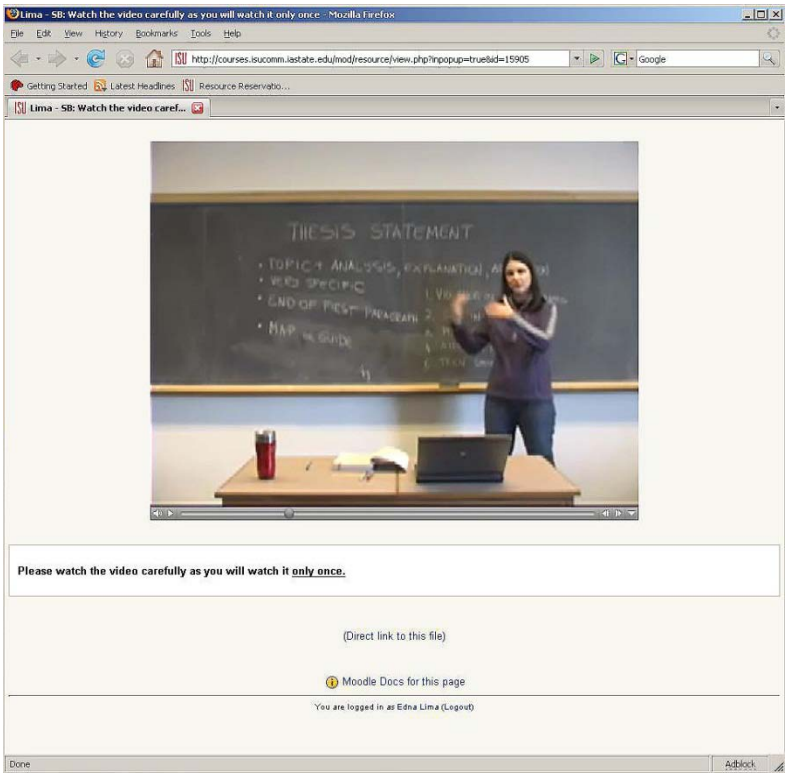


Figure 1. Screen shot of the video-taped lecture

Post-video Questionnaire

The last part of the study involved a questionnaire (Appendix A) in which participants rated the lecture and the speaker based on language and nonlanguage factors. The scale (Munro & Derwing, 1995; 1998) for each dimension ranged from 1 to 9, lowest numbers representing positive ratings and highest numbers representing negative ratings (e.g. very strong accent, very fast).

Procedures

This study was conducted during regular class periods of five sections of English 150 and 250. The data collection took place online over three different days, when the sections met in a computer lab, for about 35 minutes in each section.

Raters received step-by-step instructions on how to access the materials online. First, they filled out the personal information questionnaire. They were then directed to read the information about the speaker, watch the lecture carefully as they were supposed to watch it only once, and fill out the post-video questionnaire right after watching the lecture. Raters watched the lecture on individual computers, and headphones were used to eliminate noise interference.

Since there were five sections, the speaker to be rated in each section was alternated. Two different sections rated the Egyptian speaker, two different sections rated the Brazilian speaker, and only one section rated the American speaker. For data analysis, the two sections were combined into one to compose three treatment groups.

Analysis

In order to answer the two research questions, the eight response variables (accent, comprehensibility, speed, level of interest in the lecture, usefulness of the lecture, likeability of the speaker, teaching ability of the speaker, and teaching style of the speaker) and four predictor variables (attributed nationality of the TA, raters' gender, raters' L1, and number of ITAs) were used.

Research question one, which examines if the three groups rate the lecture differently depending on what they are told about the speaker, was answered through an ANOVA type III test with fixed effects for treatment. Given the unequal size of the samples, the least square means for each response variable were calculated. All responses were log transformed in order to stabilize the variances across treatments. The probability value (p-value) was set at .05 for all the response variables.

Question two (How do the ratings differ across groups depending on raters' gender, first language (L1), and number of international teaching assistants they had class with before the study?) was addressed by an analysis of variance type III sum of squares (ANOVA) calculations. The model designed for analysis was a generalized linear model with fixed effects for gender, first language (L1), number of teaching assistants, and TAs' attributed country of origin. The probability value (p-value) was set at $p < .05$. Since the analysis involved multiple comparisons with unbalanced design, the Tukey-Kramer method was used. In this analysis, the responses were also log transformed.

RESULTS AND DISCUSSION

RQ1: Do the three groups rate the lecture differently depending on what they are told about the speaker? If so, in what specific areas do the groups rate the lecture differently?

In order to interpret the results, it is important to take into account the rating scale used. The scale for each dimension ranged from 1 to 9, lower numbers representing positive ratings, and higher numbers representing negative ratings (e.g. very strong accent, very fast, very difficult to understand). Therefore, low means represent more positive evaluations, and high means represent more negative evaluations of the speaker. Table 2 displays the least square means found for each of the eight response variables across treatment groups and the *p* value (set at .05) for treatment group.

Table 2. *ANOVA Results across Treatment Groups*

	Response variable	Group I (Egyptian speaker)	Group II (Brazilian speaker)	Group III (American speaker)	**p value**
Language factors	*Accent*	1.33	1.33	1.59	**.19**
	Speed	.62	.69	.71	**.91**
	Comprehensibility	.68	.66	.41	**.44**
Nonlanguage factors	*Level of interest in the lecture*	1.66	1.38	1.42	**.13**
	Usefulness of the lecture	1.32	1.01	1.11	**.15**
	Likeability of the speaker	1.33	1.02	1.17	**.09**
	Teaching ability of the speaker	1.27	1.10	1.19	**.52**
	Teaching style of the speaker	1.54	1.34	1.29	**.15**

As seen in Table 2, although there is variation among the *p* values for each response variable, no statistically significant differences were found for any of the response variables. The statistics indicate that the attributed nationality of the speaker had no influence on nonnative undergraduates' perception of ITAs for both language and nonlanguage factors.

Language wise, this finding is surprising because, based on previous empirical research findings that nonnative students tend to prefer native teachers (Lasagabaster & Sierra, 2005), we would expect participants to assign more negative ratings to the Egyptian and Brazilian TAs and more positive ratings to the American TA for *accent*, *speed*, and *comprehensibility*. However, this was not the case. Bent and Bradlow (2003) found that for nonnative listeners, the intelligibility of high proficiency speakers from the same L1 background is equal to the intelligibility of native speakers (matched interlanguage speech intelligibility benefit) and that the intelligibility of high proficiency speakers from different L1 backgrounds is equal to or greater than the intelligibility of native speakers (mismatched interlanguage speech intelligibility benefit). In this study, only two listeners who were native speakers of Arabic (one from Libya and one from Saudi Arabia) rated the speaker with the ascribed Egyptian nationality. There were no native speakers of Portuguese involved in the study. Thus, the majority of the listeners (53 out of 55) rated a TA that had a different attributed L1 background. We would then assume that Bent and Bradlow's results of mismatched interlanguage speech intelligibility benefit would be more in line with the findings in this study. It is noteworthy, however, that Bent and Bradlow's first language sample was not as diverse as the one in this study. Their study included participants from four different L1 backgrounds while this study involved participants from 10 different native languages.

When it comes to nonlanguage factors, research has shown that American undergraduates tend to react to and rate ITAs more negatively depending on the country of origin of the ITAs (Brown, 1992) or their ethnicity (Rubin, 1992; Rubin et al., 1999). The ANOVA results in this study indicate that nonnative undergraduates seem not to take those factors into account when rating TAs' *likeability*, *teaching ability*, and *teaching style*. The values of the least squares means are very close for those three dimensions for the international TAs (Egyptian and Brazilian) and the American TA.

As for *level of interest in the lecture* and *usefulness of the lecture*, the results also showed no statistical significance across groups. Therefore, raters' level of interest in the lecture and their perception of how useful the lecture was were not influenced by the TA's attributed country of origin.

RQ2: How do the ratings differ across groups depending on raters' gender, first language (L1), and number of international teaching assistants they had class with before the study?

Given that first language was an important predictor variable in this analysis, three of the languages which had only one observation (Chichewa, Japanese, and Urdu) were removed from the analysis in order to yield more consistent results. The seven languages analyzed were Arabic (5), Chinese (16), Hindi (3), Indonesian (4), Korean (5), Malay (15), and Spanish (4).

Regarding *TA's attributed nationality*, Table 3 shows that statistically significant results were found for the response variables "accent" (p .04) and "likeability of the speaker" (p .04) only. As for accent, the least square means show that participants assigned more negative ratings to the supposed American TA (1.67) than they did to the alleged Brazilian (1.22) and Egyptian (1.24) TAs. This finding is surprising because previous research findings on American undergraduates' reactions to ITAs have shown that even ITAs with high proficiency in English are negatively evaluated by those students in regards to language competence (Orth, 1982; Rubin et al., 1999). Similar results would be expected from nonnative undergraduate students, but it was not the case. The alleged ITAs received very positive ratings for accent whereas the alleged American TA received slightly more negative evaluations. One possible explanation for the positive ratings for accent assigned to the Egyptian and Brazilian TAs may be the fairly high level of English of the participants and the fluent language proficiency of the speaker. Another plausible explanation may be the fact that that nonnative listeners share knowledge of the construction of the target language and they develop common strategies when learning to produce and perceive a foreign language (Bent & Bradlow, 2003). Although there is no concrete evidence as to why the raters rated the American TA's accent more negatively, the most reasonable explanation is that, taking into account that they believed they were evaluating a native speaker of English though they were actually evaluating a nonnative speaker, they found the accent to be unfamiliar to them or different from the standard norms they have been exposed and accustomed to.

Table 3. *Results of ANOVA with Fixed Effects for Attributed Nationality of TAs*

Variables	TA's attributed nationality		
		lsmeans	**p value**
Accent	Egyptian	1.24	
	Brazilian	1.22	
	American	1.67	**.04***
Speed	Egyptian	.77	
	Brazilian	.72	
	American	.63	**.87**
Comprehensibility	Egyptian	.64	
	Brazilian	.68	
	American	.39	**.64**
Level of interest in the lecture	Egyptian	1.56	
	Brazilian	1.29	
	American	1.24	**.15**
Usefulness of the lecture	Egyptian	1.45	
	Brazilian	1.12	
	American	1.17	**.11**
Likeability of the speaker	Egyptian	1.25	
	Brazilian	.85	
	American	1.01	**.04***
Teaching ability of the speaker	Egyptian	1.21	
	Brazilian	1.11	
	American	.88	**.24**
Teaching style of the speaker	Egyptian	1.43	
	Brazilian	1.33	
	American	1.10	**.20**

(lsmeans = least square means, p = p value (set at p<.05), * = significant value)

As for speaker likeability, the ANOVA tests showed that the predictor variable "attributed nationality" did influence participants' rating of speaker likeability. The least square means (Table 3) show that participants assigned more positive ratings to the supposed Brazilian TA (.85) than they did to the alleged American (1.01) and Egyptian (1.21) TAs. This finding may be explained by how the Brazilian people and culture are stereotyped around the world. Additionally, social experiences that the raters are likely to have had with Brazilian individuals in different social and cultural contexts may have influenced this finding. Unfortunately, this study did not ask participants to explain the likeability rating that they assigned to the speaker. Hence, in order to find out exactly why the raters showed a preference for the alleged Brazilian TA, a follow-up study would have to be conducted.

When it comes to participants' gender and the number of ITAs with whom they had classes prior to this study, no statistical significances were found for any of the response variables (language and nonlanguage factors). Thus, these two predictor variables did not influence their perception of ITAs regardless of the ITA's attributed country of origin. In fact, the low values of the least square means found for "gender" and "number of ITAs" indicate that raters assigned very positive ratings to the speaker for both language response variables and nonlanguage response variables. This finding also contradicts Lasagabaster and Sierra's (2005) finding that nonnative students show a high preference for native speaker teachers, especially at the university level. If they preferred native teaching assistants over international teaching assistants, we would expect to see their preference reflected on the scores assigned to the supposed American TA.

As for first language (L1), the only statistically significant result was found for the response variable *likeability of the speaker* (*p* .00). The least square means show that Arabic speakers (.67), Hindi speakers (.88), and Spanish speakers (.53) were the ones to assign more positive ratings to the alleged Brazilian TA. Since the design of this study did not include a follow-up instrument to explain raters' choices and ratings, I can only speculate the reasons for this finding. All of the three groups of speakers who assessed the alleged Brazilian TA more positively for speaker likeability are non-Asian, and two of the three groups (Hindi and Spanish speakers) are from Indo-European languages. The language familiarity may have played a role on how these two groups rated the Brazilian TA for likeability. As for the Arabic speakers, a possible explanation would be the fact that some native listeners have a hard time understanding the accent of speakers with the same language background. A more simple explanation, however, would be the stereotyping of Brazilians around the world and possible socio-cultural experiences that participants may have had with Brazilian individuals.

CONCLUSION

Research findings on American undergraduates' perceptions of ITAs' language competence have indicated that those students tend to react negatively to ITAs even when the ITAs are highly proficient in English. The findings in this study revealed that when treatment groups (grouped according to what they were told about the speaker's nationality) were analyzed separately, the attributed nationality of the teaching assistant did not influence nonnative undergraduate' ratings of language factors (accent, speed, and comprehensibility) and nonlanguage factors (level of interest in the lecture, usefulness of the lecture, likeability of the speaker, teaching ability of the speaker, and teaching style of the speaker). However, when the predictor variables were analyzed (the whole population of participants together), the findings showed that when it comes to language response variables, "accent" was the only dimension with statistically significant results when "TA's attributed nationality" was the predictor variable. Raters assessed the supposed American TA as having a slightly stronger accent than the alleged international TAs (Brazilian and Egyptian). The most plausible explanation is that, by believing that they were rating a native speaker of English, the raters found the accent to be unfamiliar or different from the standard norms that they have been exposed to.

As for nonlanguage factors, the only response variable that showed any statistically significant results was speaker likeability. The raters favored the alleged Brazilian TA when "TA's attributed nationality" and "first language" were the predictor variables. These findings may be explained through the stereotyping of the Brazilian people and culture worldwide. Moreover,

experiences that the raters are likely to have had with Brazilian individuals in different social and cultural contexts may have influenced this finding. For all the other response variables analyzed in this study, no statistically significant results were found.

The findings in the study, discussed in light of past research findings on American undergraduates' perceptions of international teaching assistants, provide valuable insights on how both groups of students react to ITAs. The results yielded by the present study indicate that American undergraduates and nonnative undergraduates have very different perceptions of nonnative teaching assistants. Based on the findings in this study, it seems that nonnative undergraduate students feel comfortable having teaching assistants who are nonnative speakers of English provided their spoken English is very good. A possible explanation may be the exposure these students have had to foreign-accented speech. Moreover, while learning a foreign or second language, learners develop common strategies to produce and perceive a foreign language.

The number of international teaching assistants in American universities is large, and it is and will continue to increase. Given that classrooms in American higher education institutions are composed of a majority of native undergraduates, it is essential that the communication process between ITAs and students be effective. Research has revealed that American undergraduates are generally dissatisfied with their ITAs' teaching and language competence. The findings in this study indicate that the opposite is true when it comes to nonnative undergraduates. Since much research has shown that the root of the problem is the negative reaction to nonnative teaching assistants by native undergraduates, higher education institutions in America need to figure out and adopt measures that will at least minimize American undergraduate students' negative perceptions of ITAs.

ABOUT THE AUTHOR

Edna Lima is a Ph.D student in Applied Linguistics and Technology and a Teaching Assistant at Iowa State University. She teaches First-Year Composition and ESL courses. Edna has taught English as a Foreign Language for over 14 years at different language institutes, programs, and at the Federal University of Viçosa, Brazil. She has also taught ESL at the Alabama Language Institute in Gadsden, Alabama. As for research, she is interested in CALL, language testing, material design, and technology applied to language learning and teaching, especially to pronunciation instruction. Email: ednalima@iastate.edu

REFERENCES

Bent, T. & Bradlow, A. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114, 1600-1610.

Brown, K. (1992). American college student attitudes toward non-native instructors. *Multilingua*, 11, 249-265.

Damron, J. (2003). What's the problem? A new perspective on ITA communication. *The Journal of Graduate Teaching Assistant Development*, 9 (2), 81-86.

Davis, W.E. (1991). International teaching assistants and cultural differences: student evaluations of rapport, approachability, enthusiasm and fairness. *Preparing the Professoriate of Tomorrow*. Chapter 56, p.p. 446-451.

Gill, M. (1994). Accent and stereotypes: their effect on perceptions of teachers and lecture comprehension. *Journal of Applied Communication Research,* 22, 348-361.

Lasagabaster, D. & Sierra, J. M. (2005). The nativeness factor: An analysis of students' preferences. *Internationaler Technischer Literaturanzeiger,* 147-148, 21-43.

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.

Lippi-Green, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States.* London: Routledge.

Munro, M. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal,* 20, 38-51.

Munro, M. & Derwing, T. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech,* 38 (3), 289-306.

Munro, M. & Derwing, T. (1998). The effects of speaking rate on listener evaluations of native and foreign-accented speech. *Language Learning,* 48 (2), 159-182.

Munro, M., Derwing, T. & Morton, S. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition,* 28, 111-131.

Orth, J. L. (1982). University undergraduates evaluational reactions to the speech of international teaching assistants. Unpublished doctoral dissertation, University of Texas.

Rubin, D. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative

Rubin, D. & Smith, K. (1990). Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants. *International Journal of Intercultural Relations,* 14, 337-353.

Rubin, D., Ainsworth, S., Cho, E., Turk, D., & Winn, L. (1999). Are Greek letter social organizations a factor in undergraduates' perceptions of international instructors? *The International Journal of Intercultural Relations,* 23 (1), 1-12.

Appendix A. Post-Video Questionnaire

Rating the lecture and the speaker

Now that you have carefully watched and listened to the lecture, please rate the lecture and the speaker according to the aspects below. Note that the scale ranges from 1 to 9, being 1 positive rating and 9 negative rating.

1. Do you know the speaker from before? Yes or no? If yes, please explain.

2. Accent

No accent	1	2	3	4	5	6	7	8	9	Very strong accent

3. Speed

Appropriate speed	1	2	3	4	5	6	7	8	9	Very fast

4. Comprehensibility

Easy to understand	1	2	3	4	5	6	7	8	9	Very difficult to understand

5. Level of interest in the lecture

Very interesting	1	2	3	4	5	6	7	8	9	Not interesting at all

6. Usefulness of the lecture

Very useful	1	2	3	4	5	6	7	8	9	Not useful at all

7. Likeability of the speaker

Very likeable	1	2	3	4	5	6	7	8	9	Not likeable

8. Teaching ability of the speaker

Very good teacher	1	2	3	4	5	6	7	8	9	Not a good teacher

9. Teaching style of the speaker

Very engaging	1	2	3	4	5	6	7	8	9	Not engaging

Levis, J. (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*, Sept. 2010. (pp. 56-69), Ames, IA: Iowa State University.

ASSESSING SPEECH INTELLIGIBILITY: EXPERTS LISTEN TO TWO STUDENTS

John Levis, Iowa State University

In order to use intelligibility as a guideline for assessing pronunciation, we need to understand how judgments of speech intelligibility are made, and to what extent judgments are consistent across different raters. This panel discussion asked English pronunciation teaching experts from around the world to listen to the free speech and read speech of a Spanish and Korean native speaker and to evaluate what features in their English most impacted their intelligibility. The discussion touched on a variety of themes, four of which are discussed: listener specific factors, identification of features that seemed to be important, the scapegoating of foreign accent, and the consequences of using read speech and free speech in assessing intelligibility.

The 2nd Pronunciation in Second Language Learning and Teaching Conference ended with an interactive discussion about intelligibility based on recordings of the reading aloud and free speech of two students, one from South Korea and one from Colombia. A panel of expert participants at the conference were asked to listen to the recordings of the students, discuss and take questions from the audience. The panel consisted of Bertha Chela (Venezuela), Tracey Derwing (Canada), John Levis (US), Greta Muller Levis (US), Murray Munro (Canada), Marnie Reed (US), Joanna Smith (New Zealand), Brian Teaman (Japan), and Beth Zielinski (Australia). The participants were asked to participate because of their expertise in teaching pronunciation.

First, all participants (including the audience) listened to the speech of two graduate students: A native speaker of Spanish from Colombia and a native speaker of Korean from South Korea. Each of these students had previously been recorded answering a set of questions in an interview format (which we will refer to as "free speech") and in reading a passage specially constructed to help to diagnose a wide variety of pronunciation errors (Celce-Murcia, Brinton, Goodwin & Griner, 2010). We will refer to this as "read speech." The reading passage and the questions are found in the Appendix. Each of the students had previously been recorded when they had taken part in an optional pronunciation tutoring component of an ESL methods class on the teaching of oral language skills. These two students' recordings were chosen deliberately for this exercise. First, they both had large differences in their performance on the free speech and read speech portions, with the Spanish speaker being fluent and relatively easy to understand in free speech and fluent but hard to understand in read speech. The Korean speaker, on the other hand, was halting in free speech and had trouble understanding the questions, but he seemed to read aloud much more clearly. Second, they came from very different L1s which may or may not have been familiar to panelists. Finally, they both demonstrated a wide variety of pronunciation problems that could be candidates for instruction.

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 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. The panel discussion was video-recorded and transcribed, and themes were identified from the discussion. Four themes will be discussed in this paper: The listener's dilemma, features thought to affect intelligibility, the scapegoating of pronunciation and accent, and the value of reading vs. speaking freely in diagnosing pronunciation. Finally, some conclusions about using intelligibility are offered based on the panelists' discussion.

THEME 1 – THE LISTENER'S DILEMMA

It quickly became clear during the discussion that everyone was not equally effective listening to both students. All listeners face challenges when speech is not packaged in expected ways, and the greater the mismatch in expectations, the more trouble listeners had. These mismatched expectations included things as mundane as not being able to clue in to the rhythmic structure, stress patterns or vowels used by the students. They also included getting lost in trying to understand and panicking, and a lack of familiarity with a particular accent.

Several panelists talked about having trouble with the Spanish speaker because the student's speech didn't sound right. This usually took the form of comments about the rhythm or tempo of his speech, which showed up in that he did not seem to distinguish between the length of stressed and less stressed words very much. He was very fluent in both free speech and read speech, and a number of panelists found that he was hard to follow because his fluency did not match their ability to listen to him. This is how several panelists described what they thought was the source of their difficulties.

> Greta: The part in the free speech that I had trouble with and I think Beth did as well, maybe like you said Murray, there were a couple of phrases that just completely lost me. And we are guessing here that it was a lack of, that there were probably many syllables that…seemed all the same to us, and Bertha's idea that there was a lack of de-emphasizing the stressed pieces in these phrases so that the prominence did not come out.
>
> Murray: And my experience that was it was a kind of tempo problem. So maybe we're talking about more or less the same thing. And I think I noticed that more in the reading passage, though. But it seemed to me that there were sometimes when he would be sort of going in a normal rate and then all of the sudden there would be a 'bada bada baa." A rapid burst and I was just lost.
>
> Greta: Yea, that's what we thought.

Later, the same theme came back again with an additional feeling attached to it: panic. This reaction to having trouble understanding is not unusual in situations where listeners have to understand and find they cannot. In this case, the experts were asked to be up front because they were experts, and they were on the spot. They 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. This is how they expressed this feeling.

> Beth: Well, whenever I can't understand anybody and can't get any English out of it, from experience, I think it's because I can't latch on to anything and understand and quite hear any of the differences between strong and weak syllables. And so that's my underlying premise: If I can't hear any English at all, I'm not hearing any words, I can't hear any word boundaries. So I think that prominence thing, the difference between weak and strong syllables is really important. And I also felt that he was pausing but in the wrong places, and so he'd go," dadadada.. dadadada". That was my impression of it, but maybe I was just panicking because I could not understand. I was thinking, "how am I going to analyze this, I can't understand it."
>
> Bertha: That was funny. The same thing happened to me with Korean. I could not understand anything and I just switched off, you know, I was, "My God" (laughter).

To remedy this, one panelist thought that the student could just be taught how to slow down a bit, so the listener could catch up.

> Joanna: I think one more thing that we see about him. Because he was reading quite fast that when he did make a mistake, as some of us know, I was just trying to figure out what he said and I lost the next bit. If he had paused, like I said, at the end of the sentence and took a breath, it would have given us time to catch up with him.. ok.. and carry on. And that's something that is kind of quite easy to teach, even without changing any of his pronunciation, just pausing so people can catch up.

One of the audience members disagreed about the effectiveness of this remedy.

> Audience Member: Well…I'm a teacher, so, I mean, just for our speaker…I used to teach him. I know him. So I was like, "Oh, that's so good. I like it." I understood every word he said…And I remember having a conversation with him…trying to get him to slow. I mean. I tried working with him…but you know, just trying to get him to slow down. But it's easier said than done. How do you get people to slow down? I think that's what he really needs to do, you know…how do you get people to slow down?

The listener's dilemma includes one more point that is of interest. Beth, from Australia, admitted that she simply did not have much familiarity with Spanish–accented English. The North American listeners, on the other hand, found this particular accent very familiar. Similarly, Bertha, from Venezuela, found the Korean speaker's accent and errors unfamiliar, and as she said, she simply switched off. It appears that familiarity with particular accents and patterns of errors helps listeners to listen more effectively, and lack of familiarity can more easily lead to panic, switching off, or even hostility. Fortunately, it is possible to become more familiar, and it doesn't take years of practice to do so, as Derwing, Munro and Rossiter (2002) found in training social work students to interact with Vietnamese accented clients. However, the experience of these experts indicates that losing the connection to a speaker's message can result in panic or switching off, even for the most open and well-intentioned listeners. For listeners who are not so well-intentioned or open, reactions may be much more negative. Listeners are not all created equal, nor are their reactions to not being able to understand.

THEME 2 – WHAT SEEMED TO AFFECT INTELLIGIBILITY

The panel was asked several times to address the question of what impacted intelligibility the most. Judy Gilbert's image of pronunciation teaching as triage was raised to ask the panel about what they would work on if they had limited time and wanted to make the biggest difference for these students' effectiveness in speaking English. What was clear was that even the experts do not completely agree on what most impacted intelligibility.

One of the things that was agreed on had to do with the Korean speaker, who was very halting in his free speech and often did not answer the basic interview questions in a way that showed that he understood what was being asked. Using a twist on Hinofotis and Bailey's (1981) idea that there is a threshold level of pronunciation that is necessary before a speaker can be intelligible, panelists felt that the Korean speaker did not demonstrate a threshold level of general speaking or listening ability, and that working on his pronunciation would not be effective until his overall speaking and listening proficiency increased. This discussion is clearly related to Suzanne Firth's (1992) argument for a curriculum that focuses first on the big picture of general speaking and listening, what she called the Zoom Principle.

> John: Let's talk about the Korean real quickly because he's a very different case especially in the reading passage, you probably could have understood many things even without the text in front of you. So what are your feelings about his real difficulties and what you would prioritize?
>
> Joanna: Well, my first impression is obviously he does not have very much experience speaking, either producing or listening because he did not understand the questions. He was completely misunderstanding them. So. And I was very surprised at how much he was able to actually read these things. Because I thought he's not going to have a chance. But he, I think he knew a lot of these words, so he must have a lot of exposure to reading and writing. But some of his pronunciations that we noticed like the word "proNOUnciation" [aʊ] are things that just once he noticed how they're actually he said – he could probably easily say *pronunciation*, but he just doesn't know what things sound like – so my suggestion was I would probably put him on a, like a graded reader, something that is fairly simple that had audio files. You could be looking and listening at the same time, just actually get exposure to what words sound like when they're said out loud.
>
> Greta: I had a kind of similar reaction that the problem in free speech was so much not just pronunciation that the other things had to be addressed. That he couldn't.. I couldn't fairly just say, "let's work on your pronunciation", that we have to say, "how can we work on your pronunciation in addition to your grammatical and vocabulary abilities". It's just they have to be integrated together for this person.
>
> Tracey: Yeah, I agree. His comprehension was really weak and he really needed a lot of help with speaking and listening.
>
> Greta: Well his speaking in the free speech piece was even more word-by-word than in the reading, and it was just teeny-tiny bits at a time…I think in ideas and phrases rather

than one word ideas.

Brian: It reminded me of a common problem I have with my Japanese students that, well, there's a general rule that you can't change something unless it's moving, right? You can't steer a car that is standing still. And I feel like with students, the student that can't even produce, you know, sentences, more than two or three words at a time. You really can't do anything with them until they're actually getting language out there for you to work on. That would be something to, I do not know if you can call it fluency, whatever. Something, actual speech, actually speaking

Someone: His mouth moving

John: I was reminded of your [addressed to Tracey and Murray] talking about the Mandarin speakers [in your research] who are kind of in a tight-knit community, and I get the feeling that this student is also in that situation. He's married. His spoken English when he came here was not very good. So, and he's probably in a field where he can get away with not speaking very much, but we think what he obviously needs to do is speak. And I thought, well he's married so you can't tell him to get a girlfriend. You know, the idea that he would get somebody that would give him input. But I remember that he also had this sense that his company had sent him here and he had to work on his studies only and he did not have time to go out and practice.

Tracey: Well, this is a huge problem for a lot of international students because if their language levels aren't quite up to where they should be, they're not going to go and get extra practice because…everything takes longer. So studying, getting their homework done. All of that is going to take longer than for anybody else. And so the last thing they're going to do is sign up for a conversation class or go out with some of the people in their class or a social event or anything like that. They're just not going to do it because they have to get that assignment done or study for that test that's coming and so they're experiencing a double whammy.

Audience Member: Maybe in those kind of cases, what we could do is use the work to kind of work on those skills. So if he knows he's going to have a presentation, then we can work on the presentation. If he knows he's going to have to read a certain amount of material and understand it, maybe I can read that material and then explain it to him or something like that. And then in that way, we can couple the two together so that he can get the practice but still say, "alright. I can get the practice but it's not going to take all that extra time. It's not going to take away from what I'm really here to do which is to study. I mean, they're paying me to get this degree."

Tracey: I think that's a very good idea, but I think there's a certain point below which.. If your English level doesn't reach a certain threshold, even that isn't going to be sufficient. And I think it is the bigger issue actually about universities accepting people whose language levels are really at the point where there're going to be able to pass successfully. And it all comes down to money.

Greta: Getting back to the Korean, there was one example in here that struck me as "this

> guy has done just about as he can do in isolation" because he had real /p/-/f/ problem in free speech, but he had them perfectly fine in the reading. And it was like, "ok, he can make those sounds, now he needs to do it, he needs to use it". And it won't happen until he's using it.

When it came to identifying particular features of speech, rhythm and tempo were talked about several times as important, as was a related area, word stress. In an area of agreement, it was felt that the pronunciation of stress and segmentals in English simply cannot be divided. Tracey Derwing talked about this interdependence in relation to the Spanish speaker. The word stress errors and the segmental errors together were closely connected to the wrong guesses made in listening.

> Tracey: …I think in the case of the Spanish speaker, I found him to be particularly clear in the oral speech but in the reading passage all of the sudden he became very difficult to understand. And there were parts that I just had difficulty following at all. And if I hadn't had the written text there, I probably wouldn't have understood what he was saying. And he went very very quickly to try to be fluent, I think, but in there in the mix he had vowels that were off. So he said things like 'nAHtive' [ɑ] instead of 'native'. And that's just not knowing how it is pronounced in English, right? Because he has the vowel. It's just that in his own language, that's a cognate and he's mispronounced it. He's left off - as Marnie mentioned - he's left off grammatical markers in places, so he said 'influence' instead of 'influences' and 'concentrate' instead of 'concentrated'. So he's left things off here and there. He has put stress in the wrong places several places. And that's really a challenge for intelligibility, right? So things like 'aRAbic' and I do not know how he said 'accuRATE' [æ] or 'accuRAITE' [e]. - 'accuRAITE' [e] I believe he said. That's truly challenging for a listener.
>
> Murray: 'asSENT' [the pronunciation used for the word 'accent']
>
> Many: 'asSENT', Yea.

Vowels came in for special notice with the Spanish speaker. One panelist joked that the student could use a few more English vowels in his speech, while the Spanish speaking panelist pointed out that his lack of reduced vowels seemed serious, and yet another noticed that vowel errors were more evident in the read speech.

Deletions of sounds and syllables also came in for mention, sometimes connected to rhythm and sometimes related to word endings, mirroring Judy Gilbert's contention that final grammatical inflections may be particularly important for speaking and listening. The comment about deleted function words was called into question by one audience member, but it was quite noticeable that the Spanish speaker did seem to leave a lot of words and syllables out because he was speaking so quickly. The panelist also raises the suggestion that the student, in his fluency, has perhaps decided that reduced function words do not really need to be worried about and that perhaps they do not even need to be pronounced. Because he cannot hear them clearly, maybe native listeners can't hear them either?

> Greta: … in the reading that there were a lot of deletions. He deleted a lot of articles and

> things like that and I wonder if that could also be addressed with that same thing that they are de-emphasized but they're there.
>
> Audience Member: I have a question, though. To what extent this affects intelligibility? Even if he deleted articles, we know that deletion of certain function words does not affect intelligibility.
>
> Greta: They may or may not. That was a just a note. That's just one example. There were quite a lot of deletions as he was reading…that sort of run-away thing, and I wonder if that bit in the spoken language that I didn't understand. Yeah, articles may or may not be important, but it's just that sense of yes, the unstressed syllables are faster but they're not non-existent to me to keep the rhythm going.

Another panelist found final deletions most important, although she admits that this is particular interest of hers and that her judgments may be biased.

> Marnie: On these sweet little missing noun and verb endings. So those are very much on my mind. And so when I listen, I automatically listen for the presence and absence of these endings and whether or not they're mispronounced. For example, the extra syllables they don't pronounce. So potentially there's a little bias in my criticism. I was actively listening for the presence of absence of those endings…with that Spanish speaker who said "has too many place. The place are historical". So maybe I just have a bias against noun and verb endings but it's spilling over now into sounding less educated less competent than the speaker might be.

This comment about bias raises the issue of pet peeves that may or may not affect intelligibility, but still are annoying to listeners. Even those who try to listen only for what affects understanding can find themselves stuck on other issues. One of the audience members raised this in regard to another aspect of the Spanish speaker's performance.

> Audience Member: Well I was just thinking about the degree to which somebody just tests your patience.…does anybody get bothered by that rising intonation at the end? dadaaA tazaaA…And I don't know if that rising intonation at the end of sentences, every sentence. Does that bother people? Is it just me? I don't know. Maybe it does not matter.
>
> Marnie: …And so, to me it goes back to the theme of our conference of intelligibility. What about things that are simply off-putting? So I think a couple of issues have come up here, if it's not intelligible, it slows us down and then we've lost whatever else they were saying while we were figuring out the first part of what they said. And then the other things that are off-putting … which goes back I think to a question you were asking…So there's the intelligibility issue but also potentially the what's stigmatizing and what's off-putting.

Another comment that came up had to do with particular words that caused problems. There is no pattern here, but the words themselves (*can/can't*) caused enough of a problem that panelists agreed that it needed to be addressed in some way. While there may not be many words like *can/can't*, this discussion points to a more general issue: sometimes pronunciation accuracy on key vocabulary is critical, even when there is not general pattern that can be addressed.

> Joanna: One thing that actually Tracey was saying is that the big issue was actually his pronunciation of, well I would say as "cAn't" [a] but I think in American would be "can't" [æ]. But because he did not, because he - what did he say again?
>
> Murray: I heard it as "accents cAn be changed" instead of *can't*. Did others get that too? I heard exactly the opposite of what was said.
>
> Joanna: Yes, but it was kind of a heavy functional load, cause if you get that wrong, the whole thing changes. So if he knows that "can" was a weak form, then he would be able to distinguish them easier somehow, the opposite.
>
> Murray: Yeah, that's kind of unusual. Yes, that's sort of an unusual negative. Normally you'd say "I can go" [ə] for the positive form and "I can't go" [æ] for the negative without the /t/ necessarily. "I can't go." Here we've got "can't be changed" as opposed to "cAn be changed". Yea, and it's stressed in both cases. And so you really need the /t/ in this case, or a nice strong glottal stop. Something to make that distinction clear.
>
> Bertha: For me, this is a point I have to make in my teaching to make clear that *can*, the pronunciation of *can* from its negative.. it's very very difficult.

Another issue was the social consequences of pronunciation errors. Regarding the halting speech of the Korean speaker, and the fact that the student felt that his English might be good enough to get by, this discussion occurred.

> Joanna: I just want to comment also, it's interesting the fallacy that he can get by. If he's going to work in an English-speaking company, even engineers need to have social chitchats. There's a lot of research done in Wellington's Language in the Workplace Project (http://www.victoria.ac.nz/lals/lwp/) and it's all social chitchats. "I could do my engineering, but I can't for the life in me say hello or how are you to my workmates." And it affects their ability to function. So it's, they need to know that they have to.
>
> John: Jenny Miller's work in Australia (Miller, 2003) also has that same thing that kids in schools and in high school, immigrant teenagers, if they can't break in to the social milieu and be, she says, "if they can't be heard" by their peers, then they remain inaudible. And that the ability to do things like, she says, use things like 'like' and understand humor, that that allows other people to see them as real people and that then opens the opportunity to begin to define yourself and your identity and allows you then to acquire the language more successfully.

Finally, panelists addressed the problem of multiple errors and their cumulative effect on understanding. Although Munro and Derwing (2006) addressed the question of functional load and frequency in a limited way, and Zielinski (2008) showed how stress and segmental errors can compound effects on intelligibility, this is an area that is clearly ripe for further research, even though such research may be difficult to carry out.

> Marnie: I'll give one quick example of one of these if I may for a Korean speaker that occurred with my co-author. A student came up to her during office hours to talk about

> her thesis statement that she had to write and she plopped it down the teacher's desk and said: "I need to talk about my thesis statement, it is occurred." So my co-author thought: "well, I can see it has happened (laugh) what a strange turn of phrase." The Korean speaker, she was not trying to say it had occurred. She left out the glide, /w/ sound, but she put the stress in the wrong syllable. She was trying to say: "it is awkward."
>
> Tracey: And one thing I think in research that we haven't looked at very much at this stage is the interaction effect of different kinds of errors, right? So a lot of people have isolated different types of errors and have looked at whether or not they influence intelligibility or affect intelligibility one way or another. And there are a lot of really good studies that have done that. But there is very little or I'm not aware of any actually.. any research that has been able to sort of take say three different types of errors, look at those and then look at them together, and look at the interaction of those errors. And how, you know, is this an exponential kind of problem when you add one? Is it additive or is it worse than that compound?

THEME 3 – ACCENT AS SCAPEGOAT

Business writers sometimes talk about "The Peter Principle" in which people eventually rise to the level of their incompetence. The Korean speaker's graduate studies reflect this issue, in that his weakness in speaking and listening made it more difficult for him to study effectively in English. His pronunciation was clearly impaired as well, but listeners may hear the pronunciation errors first and think that he had to work on his accent. In reality, his much more serious problems had little to do with pronunciation. The tendency to attribute to pronunciation what may be a very different problem was seen first in relation to the discussion about the perceived speed of the Spanish speaker, but was also raised in relationship to vocabulary and other errors that were evident in the free speech portions.

> Audience Member: And also intelligibility is not only to pronunciation alone, and also collocation use because sometimes we have some predictions what kind of words .. students will say, a native speaker would say.. come up with an awkward combination and catch our ears or interfere with intelligibility.
>
> John: Yeah. 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 a cultural level, you can impair intelligibility. And all of those things need to be addressed. It just seems that in the literature, and you can help me with this, that pronunciation shows up as impairing intelligibility much more. And I think it's not because it impairs it much more, but it's the first step. And then if somebody doesn't understand…they start to grab for causes and they say, "it's an accent" or "they're not pronouncing right" or something like that. But it may be more pervasive in the kinds of causes, and the listener is not seeing it.

This response led to a discussion of accent as a scapegoat for other problems, with the example being given of presentation and teaching skills, an important issue in the training of international teaching assistants.

> Tracey: Absolutely. I think accent is the real scapegoat because it's salient, it's noticeable and so people blame it on that. I've been involved in a couple of research projects in Alberta where we were looking at international teaching assistants. And in one of them we worked with these people who'd been identified as having "terrible accents" by their students. And they had, they did have a terrible teaching evaluations, right? And so they brought their teaching evaluations to us as a team, and we worked with them. We video-taped them at the beginning of the term and then taped them at the end of the term and we worked with them all term long. And, you know, each of these individuals had an accent but actually that was not really the problem for most of them. Most of them really weren't very good teachers. And just didn't know what to do, they did not know where to start. So we spent a lot of time on presentation skills and pedagogical skills. And at the end there was a lot of improvement, but their accent hadn't changed. And one of those guys got a teaching award at the end. And he was, you know, he was incredible, like the difference was night and day, his pre- and post-video. You know, he was just so much better because he had a better idea of how to present what he was teaching to his students. And so, you know, his students had all blamed his accent. His accent didn't change one iota over the course of that term, but all of his presentation skills had changed, and he was a lot more approachable, and he, you know, worked directly with the students, he gave examples, he had handouts in advance. He went through concepts in advance instead of talking to the board which .. before.. you know. And so the students have a hard time discerning what's really the problem, but they can hear the accent, so they just blame it on the accent. But it isn't necessarily the accent.

THEME 4 – ASSESSING WITH FREE SPEECH VS. READING ALOUD

We most often assess or diagnose pronunciation through the use of a written passage. This method has a long history in pronunciation teaching, and it has many advantages. Teachers can carefully construct a single passage to include and adequate representation of all the pronunciation features that are considered to be important for instruction, and other elements like vocabulary choice or grammar can be controlled.

Nevertheless, reading aloud is not speech, and it is a task that is full of pitfalls. The way that the two students negotiated reading aloud and free speech was the subject of an extended discussion regarding how we should evaluate intelligibility. Several assertions about reading aloud were made: that it is a strange activity for most people who never have to read aloud, that it is neither a reading skill nor a speaking skill, that it promotes worse performance than free speech for some people but not for others, as well as raising questions on what the best way to use reading aloud might be. Free speech, on the other hand, was implicitly thought to be better for assessment, but this belief was not discussed in explicit detail. The main discussion centered on the problems with reading aloud.

> John: I'm curious what you think, so much of what we do in pronunciation involves reading aloud. And with this Spanish speaker, a lot of the problem is that he is not taking the written text and decoding it into a spoken form that is recognizable. How important is it with especially with this kind of level of student at a graduate school level to work on that kind of decoding? You know, the connecting of something that you can turn written speech into spoken speech. Wayne Dickerson is not here, but it's something that he has

argued for many years. And he [the student] is a good example of someone who can't do it.

Bertha: Well, as I tell many of my teachers, not to use reading aloud in class unless you've prepared the students to read. You know, you tell them where to chunk, where the stresses are, how you pronounce words, then you can read aloud. But judging students from read-alouds is very bad. It's, they're concentrating on something else, you know. And then studies of broadcasters in Spanish, Spanish broadcasters, that read the news and they go berserk, "ok, oh my God, where did I get that?" So I say it's worse for the EFL students.

Murray: I think asking people to read aloud is asking them to something that's kind of weird. It's not something that most of us are called on to do on a regular basis. I mean, it's one thing if we were reading to our children or something like that, but most of the time, it is not something that we do a lot of, most people, now some people do. And so this is an alien thing. And when you're asked to do an alien thing, it comes out weird because you are not focusing on the meaning, you're not focusing on the processing, what's being said. And so you pause in the wrong places, and you have reading pronunciations, where you pronounce something simply the way it's spelled as opposed to the way you have already learned that it's actually produced because you're not making the connection between your internalized knowledge and what's on the page. So I really agree with you.

Bertha: I said this to my teachers. One of my teachers was dictating something to a group of students and she said.. she said, "she celebrated it on Thursday'. So she said, 'did you understand?" She said it again, "she celebrated it on Thursday." She realized they didn't understand so she actually said, she said," she celebrated IT on Thursday." (laugh) So everyone was pleased. And she said it three times after that, "she celebrated IT on Thursday". I said," oh my God. The students will go out sounding like this." And this was a teacher, you know, not a student.

Tracey: I think even if you have a native speaker or a group of native speakers and ask them to read aloud, you'll find that most of them don't do a terribly good job in parsing and, especially for evaluation (unclear). And I think it's an unnatural task unless you've practiced it and unless you're going to be reading the news or an actor or something like that then most people don't have a real need for that.

Marnie: So in our master's program…we have a guest speaker who comes in. She's the author of "Short Course on Teaching Reading Skills", and then she's also written "Reading Power", "More Reading Power", "Reading Power for the Beginners" (laugh) and all of these things. And she lists reading skills, she's got a list of 24 reading skills, and reading aloud is not on it, but she'll talk about it as a good diagnostic tool. So it might give us a first pass we are exploring some of the potential problems for our students, but not to consider it a reading skill. And a good example about what would be, even for native speaker to read something aloud and then you asked a comprehension question, "I don't know, I wasn't listening to what I was saying" (laughter).

One audience member suggested that reading aloud was only valuable in certain contexts.

> Audience Member: Wouldn't it really depend on the context that he will use English in? So let's say he's never going to read aloud, and his free speech sounds ok. I mean, I don't know. Again, I think that just if you just go by reading, you're kind of putting him at a disadvantage, because, you know, if I had to read Spanish, I'm sure that I, the same problems that he did. But if my fluency in spoken, in just regular conversational speech is fine and that's the goal, then why mess with it? You know…if he can be intelligible in conversational speech, then why continue to beat a dead horse, you know?

Not explicitly addressed in this discussion of reading aloud is whether teachers can effectively make use of written text and predictive rules, as argued by Wayne Dickerson (1987). What seems to be clear is that using written texts to assess intelligibility is problematic. Reading aloud is not the same as speaking, and unless students are given sufficient preparation in the task, the task itself is likely to create errors that do not show up in free speech. However, it is also clear that in academic settings, students need to learn how to better connect spelling patterns to what the words actually sound like.

CONCLUSION

A number of conclusions are suggested from the discussion. First is that intelligibility is a moving target, depending on the type of listening being done, the familiarity of the listener with the patterns of the speaker, the, the types of language features that cause conflicts with the expectations that a listener brings to the interaction, and the type of speaking task. In addition, intelligibility may be seriously impaired by a single error (as with *can/can't*) or it may only be impaired when multiple errors combine to violate expectations to the extent that listeners can't process speech successfully in real time (as in *occurred vs. awkward*). In addition, although pronunciation may be frequently implicated in loss of intelligibility, pronunciation is clearly not the only issue. It may be the most noticeable, but pronunciation never exists on its own in communication, and vocabulary choice, grammar, fluency, listener internal factor, speaking choices, and the information presented through other channels (such as the use of the board in a lecture) all may contribute to a loss of understanding.

ACKNOWLEDGMENTS

Thank you to the panelists for making this discussion possible, and to Ghinwa Alameen for her work in transcribing the video.

ABOUT THE AUTHOR

John Levis is Associate Professor of TESL and Applied Linguistics at Iowa State University. He teaches courses in ESL teaching methods, the teaching of pronunciation, speaking and listening, Introduction to linguistics, Sociolinguistics, and the history of English. He is the co-organizer of the Pronunciation in Second Language Learning and Teaching Conference. His articles have appeared in *TEOSL Quarterly, World Englishes, Language Awareness, Applied Linguistics, The Annual Review of Applied Linguistics, ELT Journal*, *SPEAK Out*, *System* and *TESOL Journal*.

REFERENCES

Celce-Murcia, M., Brinton, D., Goodwin, J. & Griner, B. (2010). *Teaching pronunciation: A reference for teachers of English.* Cambridge: Cambridge University Press.

Derwing, T., Munro, M. & Rossiter, M. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingualism and Multicultural Development, 23,* 245-259.

Dickerson, W. (1987) Orthography as a pronunciation resource. *World Englishes, 6,*1, 11–20.

Firth, S. (1992). Pronunciation syllabus design: A question of focus. In S. Avery & P. Ehrlich, *Teaching American English pronunciation,* Oxford: Oxford University Press, pp. 173-183.

Hinofotis, F. & Bailey, K. (1981). American undergraduates' reactions to the communication skills of foreign teaching assistants, in J. Fisher, M. Clarke, and J. Schachter, (eds). *On TESOL '80 : Building bridges : research and practice in teaching English as a second language : selected papers from the fourteenth annual convention of Teachers of English to Speakers of Other Languages, (pp. 120-136).* Washington, D.C.: TESOL.

Miller, J. (2003). *Audible difference: ESL and social identity in schools.* Clevedon: Multilingual Matters.

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

Zielinski, B. (2008). The listener: No longer the silent partner in reduced intelligibility. S*ystem 36,*1: 69-84.

Appendix

Free speech interview questions

1. What is your name and where are you from?

2. Tell me something interesting about your hometown.

3. What are you studying at ISU? Can you tell me a little about it?

4. If you could travel anywhere in the world, where would you go? Why?

Read speech reading passage (from Celce-Murcia et al. 2010, p. 481)

Is English your native language? If not, your foreign accent may show people that you come from another country. Why is it difficult to speak a foreign language without an accent? There are a couple of answers to this question. First, age is an important factor in learning to pronounce. We know that young children can learn a second language with perfect pronunciation. We also know that older learners usually have an accent, though some older individuals also have learned to speak without an accent.

Another factor that influences your pronunciation is your first language. English speakers can, for example, recognize people from France by their French accents. They can also identify Spanish or Arabic speakers over the telephone, just by listening carefully to them. Does this mean that accents can't be changed? Not at all! But you can't change your pronunciation without a lot of hard work. In the end, improving appears to be a combination of three things: concentrated hard work, a good ear, and a strong ambition to sound like a native speaker.

You also need accurate information about English sounds, effective strategies for practice, lots of exposure to spoken English, and patience. Will you make progress, or will you give up? Only time will tell, I'm afraid. But it's your decision. You can improve! Good luck, and don't forget to work hard.

Dickerson, W. B. (2011). Upstream destressing: Another step toward natural speech. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (pp. 70-81), Ames, IA: Iowa State University.

UPSTREAM DESTRESSING: ANOTHER STEP TOWARD NATURAL SPEECH

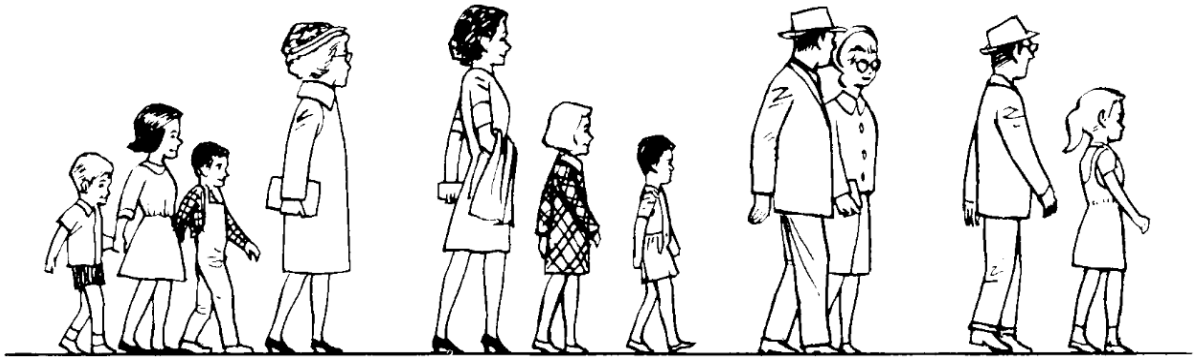
Wayne B. Dickerson, University of Illinois at Urbana-Champaign

English speech is known for its bumpiness—peaks and valleys of stress with one dominant peak variously called the discourse stress, primary phrase stress, sentence stress, nuclear stress, and tonic.

Less well known is the fact that native speakers selectively demote the peaks adjacent to the discourse stress to valley status for a variety of reasons. This is the phenomenon of destressing (or deaccenting).

This paper focuses on the least well-known part of the destressing story—upstream destressing. It is a systematic downgrading of the peak or peaks just before the discourse stress. We first describe what is involved phonetically in destressing, why this phenomenon occurs, and the rules governing its use. Then we examine the pedagogical implications of this phenomenon: How can we introduce the destressing behavior to linguistically naive learners? How can we help them predict when to destress peaks on their own? What does a destressed peak sound like? What kind of practice do they need so that this behavior will become a natural part of their oral performance?

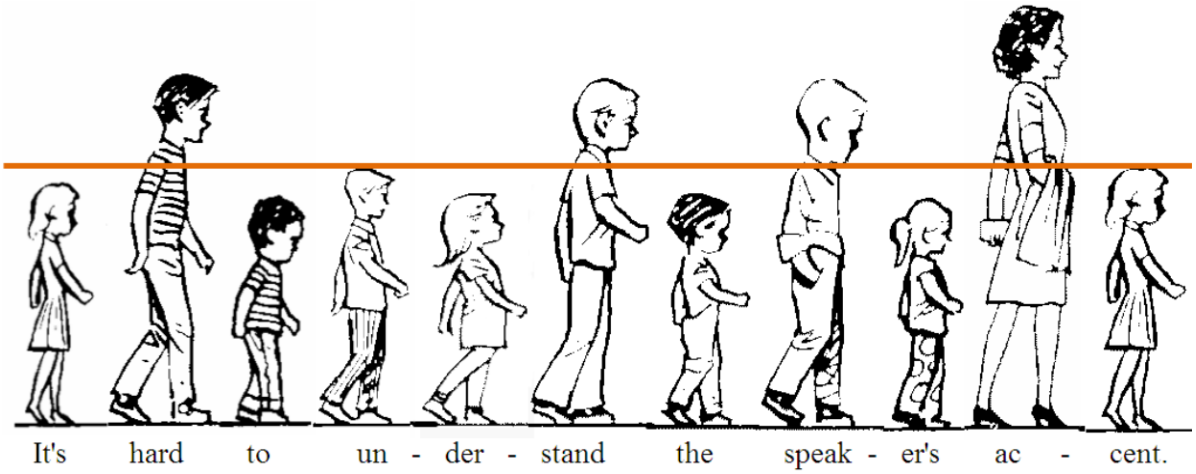
This paper centers on a stress behavior that is missing from the usual presentation of English rhythm and discourse stress in ESL pronunciation texts. The question is whether or not it should stay missing. As background, let's start with what is generally presented to ESL students about English rhythm and discourse stress. We begin with this famous picture of English rhythm from Prator & Robinett (1985, p. 29).



When the images of people are mapped on to a sentence like, *It's hard to understand the speaker's accent*, the graphic is sufficient to remind us of the key points.

First of all, English rhythm is really bumpy. We can describe this bumpiness using the landscape metaphor of peaks and valleys, by which we mean that some syllables (those above the line) have significantly greater stress—are louder, longer, and often higher pitched—than other

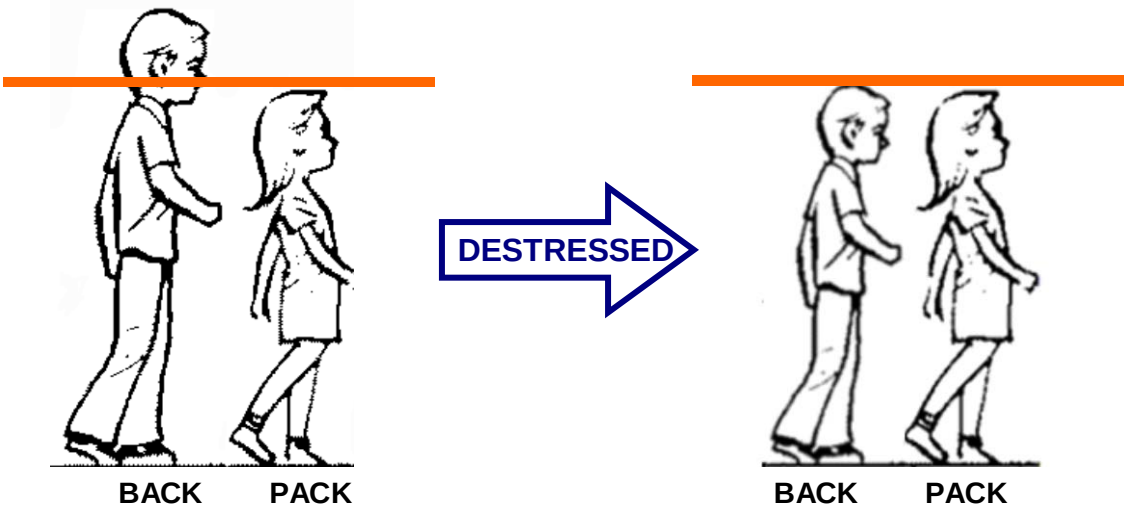
syllables (those below the line).



Second, each peak is the major stress of a content word. The syllables that populate the valleys are the non-major stressed vowels of content words and the vowels in function words—all below the line.

Third, a refinement distinguishes loud function words from soft function words. Loud function words are demonstrative pronouns, question words, and negative words and negative contractions—all of which behave like content words in that they carry a rhythmic peak. By contrast, soft function words—like conjunctions, articles, prepositions, pronouns, auxiliaries, modals, forms of the verb *to be*—belong in valleys.

Finally, although less commonly, pronunciation texts point out that any peak (content word or loud function word) coming after the primary or the nuclear stress in a phrase will be destressed or deaccented. That is, it will lose its rhythmic beat. One reason to destress words in that position is that those words contain information already mentioned in the discourse or assumed to be shared by the participants. They no longer need the listener's greatest attention. As an example, let's say that we've been talking about backpacks. In this context, I say, *I've already bóught her a backpack. Backpack*, which would carry a rhythmic peak on *back* if it were new information, is now destressed following the nucleus on *bought*. *Backpack* becomes part of a valley.

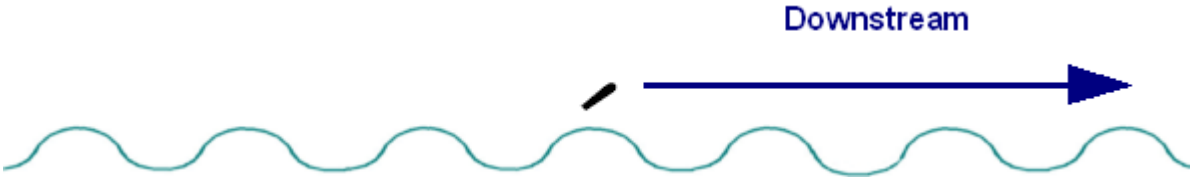


Since destressing is our topic, it is important to understand what this term refers to. First, destressing does not mean making a stressed vowel unstressed; when we destress the *back* of *backpack*, *back* does not become unstressed. Second, destressing does not affect vowel quality; it does not make a full vowel into a reduced vowel. The /æ/ in *back* of *backpack* is still /æ/ after being destressed; it does not become schwa. Third, destressing does not affect all stressed, full vowels. In *backpack*, both vowels are stressed and full, yet *pack* is unaffected by destressing. Destressing affects only the stressed, full vowel in the one syllable of a content word or loud function word that would ordinarily carry a rhythmic beat. In the case of *backpack*, that's the vowel in *back*. Fourth, the effect of destressing that particular vowel is to lower its stress so that it no longer carries a beat. Such stress lowering puts *back* into a rhythmic valley with *pack*.

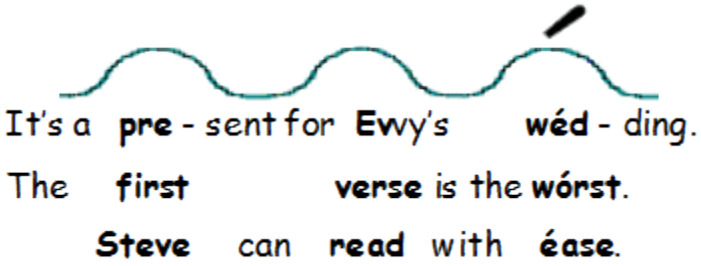
These basics about English rhythm and discourse stress are present in one way or another in many ESL pronunciation texts (e.g. Beisbier, 1995; Hahn & Dickerson, 1999; Prator & Robinett, 1985).

To go beyond this basic description, it will help to shift images. Let's borrow and extend the metaphor of the stream of speech, where the peaks in this stream represent content words and loud function words. The primary stress is one of these peaks.

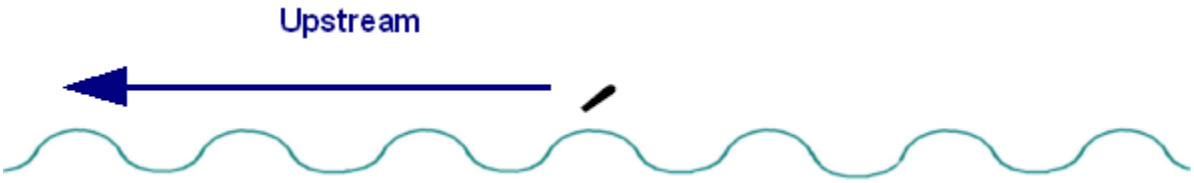
Continuing this metaphor, we can say that the destressing of *backpack* in *I've already bóught her a backpack* happens downstream from the primary stress—after the primary stress. Post-primary destressing is what we call **downstream destressing**. Bing (1980) and Bolinger (1986) discuss this notion at length because of its meaningfulness in discourse.



If we use this model for sentences with three peaks, they should easily fit into this diagram, contributing two rhythmic peaks up to the primary stress, as shown below.

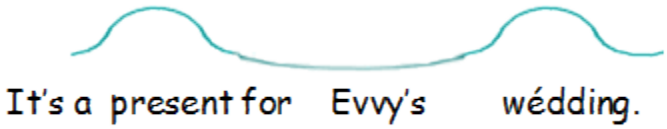


Here is where we leave the well-trodden path. What is much less well known and not represented in any ESL pronunciation texts I am acquainted with, is that the peak - valley alternation illustrated above is in fact quite rare in ordinary speech. That's because of another destressing phenomenon that happens before the nucleus, upstream from the primary stress. It's what we call **upstream destressing**.



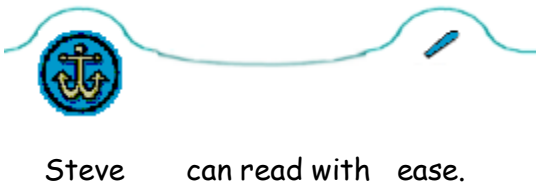
Since upstream destressing is the more specific focus of this paper and is not so well known, let's find out a bit more about it. Where does it occur?

First, like downstream destressing, upstream destressing affects only content words and loud function words—the kind of words that typically form rhythmic peaks. The effect of upstream destressing is to push these peaks into a valley, removing their rhythmic beat. In the following example, the first syllable of *Evvy* joins the preceding valley syllable, *for*, and the following valley syllable, *-vy*, to create an enlarged valley.



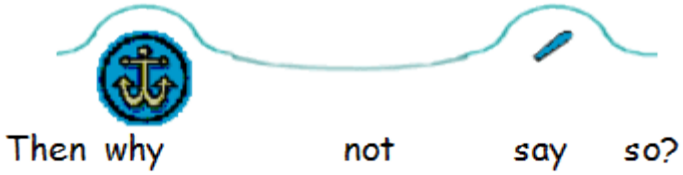
Second, it happens most commonly to the content word or loud function word nearest the primary stress on the left. It may happen further upstream, but only if it also happens nearest the primary stress.

Third, it doesn't happen if a content word or loud function word left of the primary stress is the first rhythmic beat of the phrase. So it wouldn't happen to *Steve* in *Steve can read*. But in the case of *Steve can read with ease*, it will happen—to *read*, right before the primary stress. The first rhythmic beat of a phrase acts as a kind of anchor in the phrase; that's *Steve*. In British parlance, *Steve* is called the onset of the phrase (Wells, 2006, p. 8). A peak stress in the anchor position is quite stable; it's not demoted.

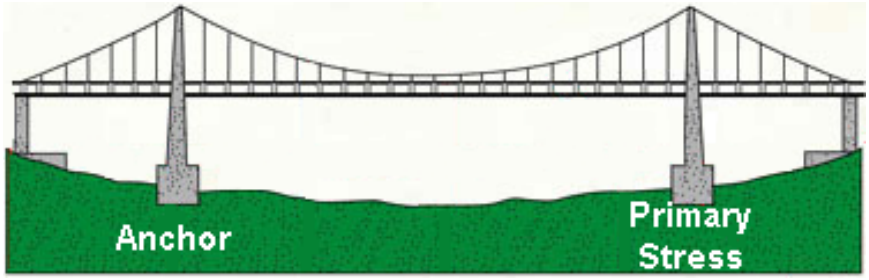


In short, the minimum conditions for upstream destressing to occur are three rhythmic stresses—an onset to the phrase, a primary stress, and a rhythmic stress between the two. It is this middle one that is ordinarily demoted to valley status.

Does the kind of rhythmic word between the anchor and the primary stress make a difference? Apparently not. Destressing seems to happen whether the middle peak is a noun, an adjective, a verb, or an adverb—a content word or a loud function word (as in the next example). *Why* is the anchor; the primary stress is on *say; not* is demoted.

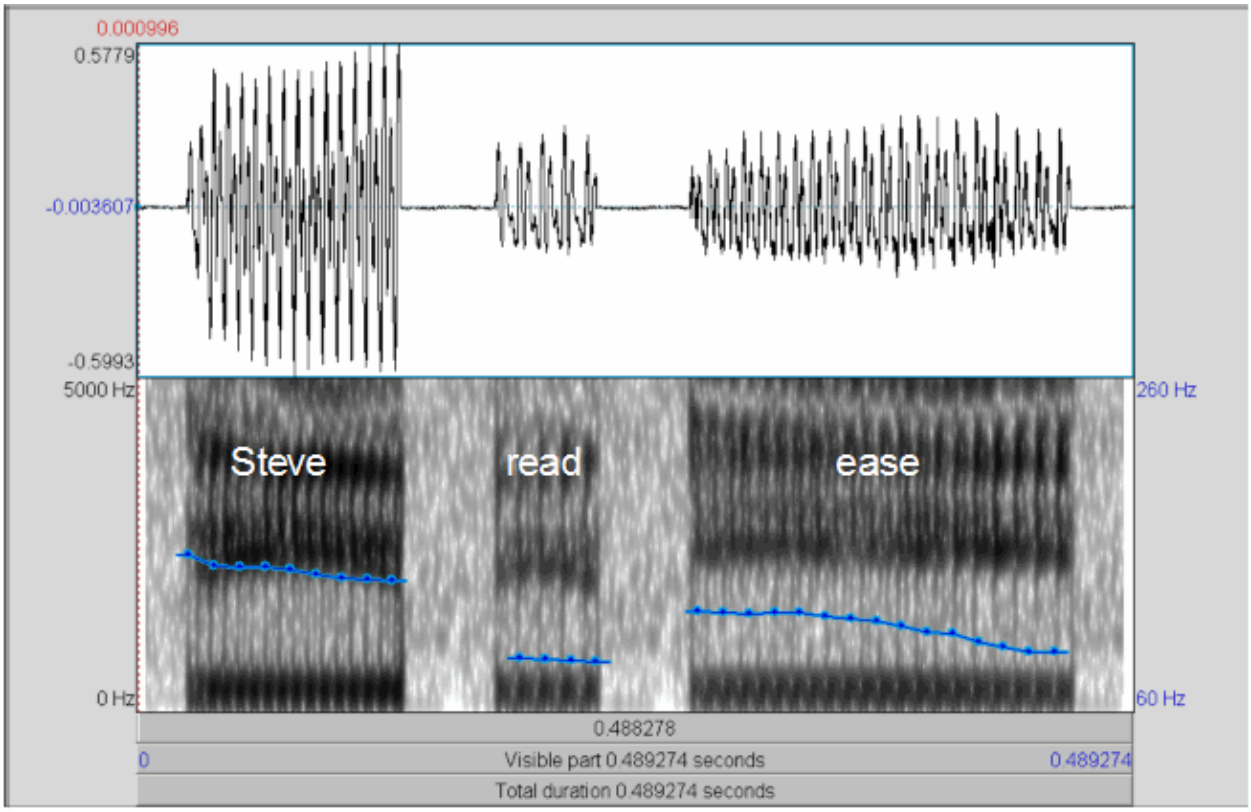


Referring to this overall pattern of anchor and primary stress ('positions of greatest impact'), Bolinger (1986, pp. 47-48) calls it a **suspension bridge**, where the 'two-accent shape' forms the 'two towers of the bridge'. Wong (1987, p. 56) picks up this apt metaphor in her guide to ESL pronunciation teaching, focusing on the towers without mentioning the destressing phenomenon between. Bolinger refers to the presence or absence of 'the sag in the suspension bridge' only as 'normal variants' (p. 47). Even though they do not give attention to what we call upstream destressing, Bolinger's metaphor is appropriate, as the graphics above and below illustrate.



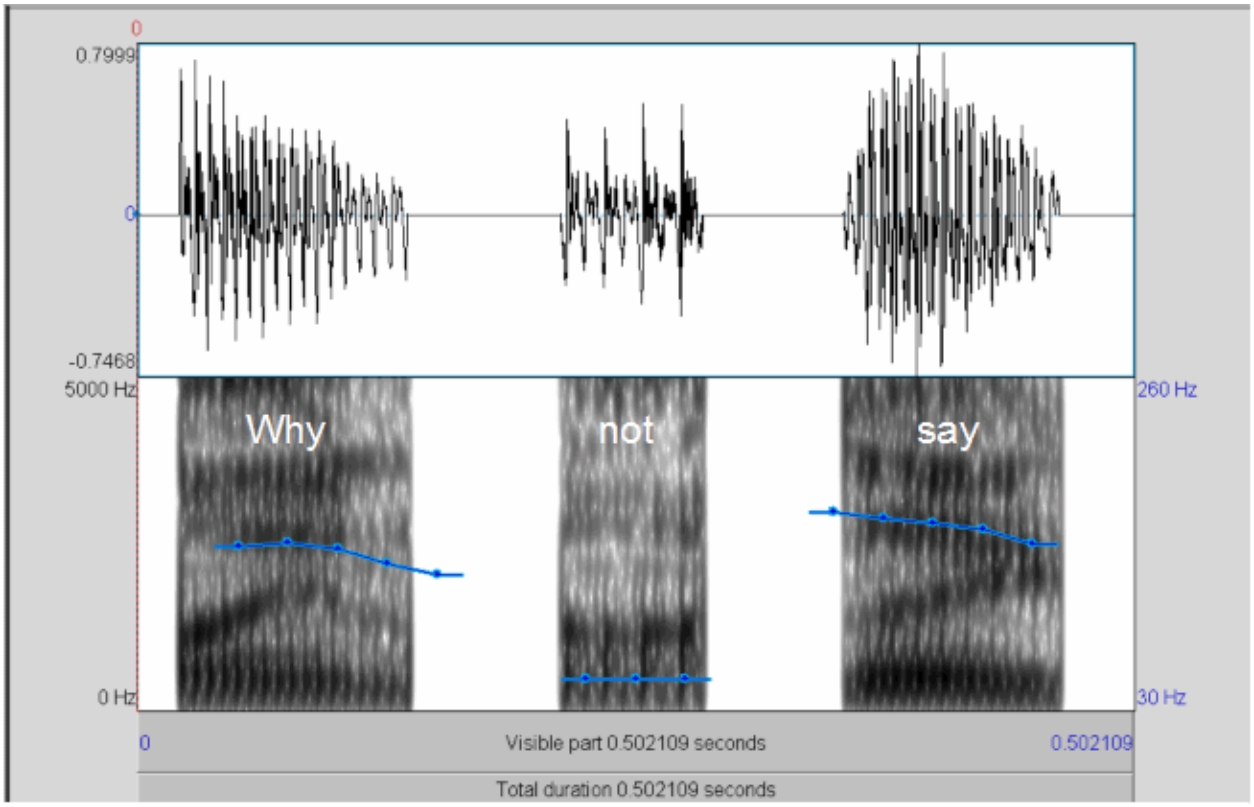
In contrast, Wells (2006) does take an interest in upstream destressing, saying, "The option to downgrade potential accents is a pervasive characteristic of English rhythm. It tends to operate whenever an accent is located between two other accents in the same IP [intonation phrase]" (pp. 229). Rather than appeal to a metaphor, Wells frames his discussion in terms of a rule he calls the Rule of Three.

What happens phonetically when we destress? As we know, stress in English is some composite of duration, pitch, and intensity. The consistent effect of destressing is to shorten the *duration* of the vowel. The following graphic displays the three [iy] vowels in *Steve can read with ease,* where the primary stress is on *ease*. With time along the bottom axis, the vowel in *read* is obviously much shorter than the vowel in *Steve*.



Depending on the intonation pattern, destressing may also drop the *pitch* of the word between the onset and nucleus as indicated by the blue lines at the bottom of the image. Furthermore, destressing reduces the *intensity* or loudness of the vowel. That is evident in the height of the wave form. There is much less energy in the [iy] of *read* than in the [iy] of *Steve*.

The same pattern is evident in *Why not say so?* where loud function words occupy the anchor and the position before the primary stress.



To understand the role of destressing, it may be helpful to compare and contrast upstream destressing with downstream destressing. While both are similar in affecting rhythmic peaks, they have quite different, non-overlapping functions otherwise. Upstream destressing has nothing *specific* to say about the message content. This is no doubt why Bolinger, who focuses on meaningful signals, dismisses the 'sag in the suspension bridge' as one of several 'normal variants' (p. 47). However, when it comes to enhancing the perception of the primary stress, upstream destressing trumps all. It serves as a marker or a listener alert: *Here comes a primary stress! Listen up!* It anticipates the upcoming nucleus by dropping the voice into a valley before the final peak. Furthermore, by expanding the valley before the primary stress, it maximizes the contrast between this valley and the primary peak so that the primary stress is strikingly highlighted.

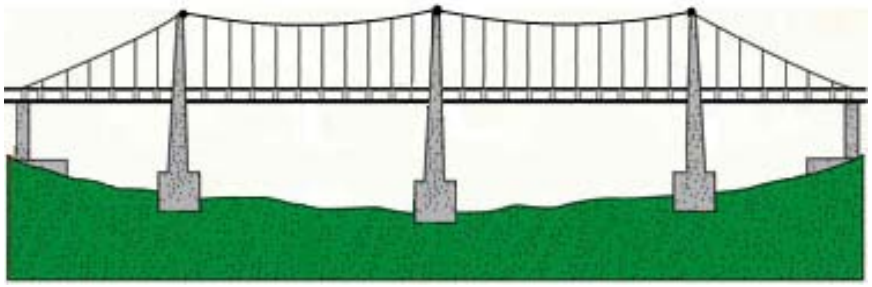
By contrast, downstream destressing, coming as it does after the primary stress, is not capable of serving as an alert. But as for carrying specific information about the relative importance of message content, downstream destressing is unsurpassed in this role. In no uncertain terms, its quiet, backgrounded delivery tells the listener: *We're not central. Focus elsewhere* (Bing, 1980, pp. 173-208; Bolinger, 1986, pp. 110-135).

Downstream destressing is clearly important to learners because of its message-bearing capabilities. But what about upstream destressing? Is there value in its alerting function and its role of making the primary stress stand out?

Let's put this question in pedagogical perspective. Pronunciation teachers know that it takes considerable effort to get students to make a remarkable primary stress. Our typical strategy is to

focus on the nuclear vowel. We emphasize a significant change in its duration, a marked change in its loudness, and a sudden change in its pitch. The approach is all about the nuclear vowel.

The problem with this approach is that students often don't make the primary stress stand out any more than other rhythmic peaks in the phrase, so there is nothing distinctive about any of the 'towers of the suspension bridge'. Or they tend to elevate the prominence of all the peaks, like making all the towers taller.



But what if one of the important ways to make a primary stress noticeable is to increase the contrast between the valley right before this major peak and the peak itself? That is, instead of focusing only on the character of the primary stressed vowel, we might help learners even more if we also show them how to demote the peak just before the primary stress. It may take such a double-barrel attack to help learners make the nucleus as prominent as native speakers do.

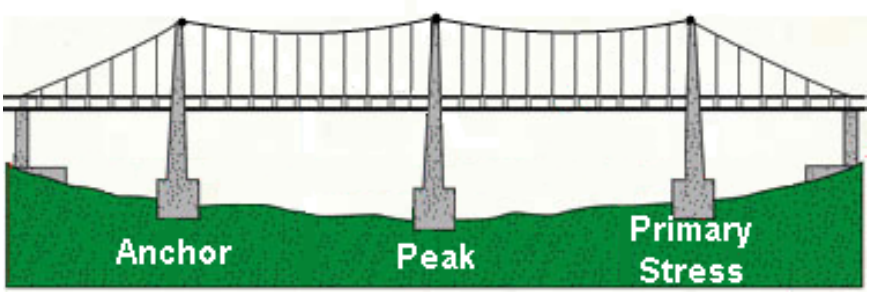
Furthermore, by giving attention to upstream destressing, students will not only foreground the primary stress and alert the listener to its imminent appearance, but their destressing will also deliberately remove potential competitors for attention—other peaks that could misdirect a listener's focus away from the central message.

Is there any evidence that the speech of learners is more intelligible and less distracting to listeners when learners demote peaks before the primary stress than when they don't? This is an important research question that awaits an answer. Anecdotally, however, when students in my pronunciation class do demote peaks as described, their speech sounds surprisingly like the English we are targeting.

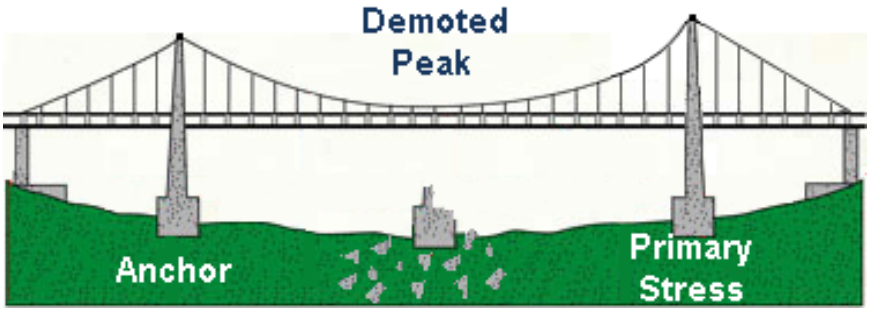
Until we have more definitive evidence, let us assume that there could be pedagogical value in teaching upstream destressing. How might this notion be introduced, predicted, pronounced, and practiced when the audience consists largely of linguistically unsophisticated learners of

English? As we shall see, the challenge is not so great as it might seem.

Introducing demoted peaks. Adult learners want to understand what we are trying to teach them and why it is important. To convey the concept of destressing, we use the graphic of the suspension bridge and identify the critical requirements for destressing: the anchor, the primary stress, and the intervening peak.

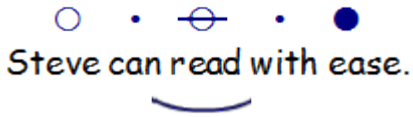


The objective of our practice is to make the primary stress stand out so that it communicates clearly to the listener. We do that by creating an extreme contrast between the valley before the primary stress and the primary stress itself. So we attack two vowels—the vowel of the peak before the primary stress and the vowel of the primary stress. The first, we make less prominent; the second, we make more prominent.



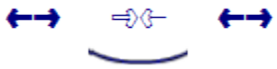
Predicting demoted peaks. To empower adult learners, we also show them how to predict upstream destressing on their own. We already help them predict rhythm by using big circles for peaks and dots for valleys, and we help them predict the location of the primary stress which we mark with a filled-in circle.

Instead of stopping here, we ask students to take another pass through the sentence to see if any upstream destressing is present. Using short sentences like the one above, we have students practice identifying the critical elements: Where is the primary stress? (*ease*) Is there an anchor? If so (*Steve*), is there a peak between the anchor and the primary stress? If so (*read*), what should we do with it? It gets a valley mark ‿. Learners become very good at determining where they should destress a peak.



Pronouncing demoted peaks. Once students find a peak to demote, how should they pronounce it? How do they make the middle tower crumble? The answer, of course, is to do the opposite of what we do with the primary stress. Using phrase sets, all with the same rhythm, we focus on each of the three dimensions of stress. (If we change the rhythm pattern with every practice item, learners will not get the feel for what is happening so easily.)

First, we focus on vowel duration. Judy Gilbert (2005, p. 18ff) hit on the technique of using rubber bands between the thumbs to have students feel the duration dimension. This kind of proprioceptive tool helps learners connect the stretching of the band with the lengthening of peaks and the relaxing of the band with the compression of valleys.



Betsy stayed at home

Try to come at five

Second, we attend to the intensity dimension. The technique for this dimension is more challenging to do, but it directly addresses the energy contrast. We simply whisper valley syllables. We start slowly and build up.

Lunch is served at noon

Bill's good at math

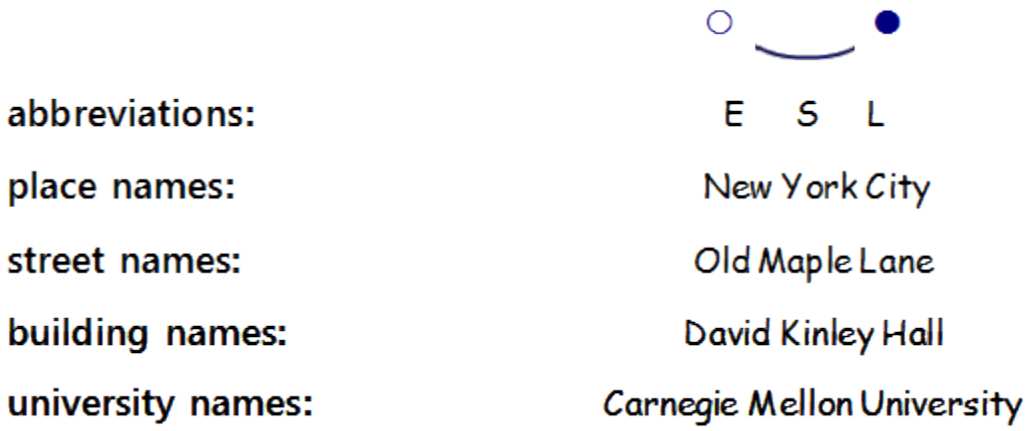
Third, the technique for the pitch dimension is to hum the word pitch, using a high pitch on the peaks and a low pitch on the valleys.

Rice is good with beans

Why didn't he come?

Practicing demoted peaks. Once students can predict where destressing should occur and what it should sound like, the principal challenge becomes giving them enough practice that demoted peaks begin to sound right to them. If Wells (2006, 229) is right, and upstream destressing is pervasive, then it takes no special materials to practice it on an on-going basis. Whenever we

give attention to the primary stress of phrases and sentences, we use the occasion to ask students to check for demoted peaks. For example, when we work on compound nouns that have phrasal stress, we routinely ask: Are there any demoted peaks? Since they can spot the anchor and predict the heavy stress of the construction, it is easy for them to locate an intervening peak to demote. Furthermore, when we give them opportunities to record sentences and phrases with demoted peaks, their practice begins to sound much more authentic than when they do not take this simple, extra step.



In summary, upstream destressing is a widely attested phenomenon that English speakers appear to use to highlight an upcoming primary stress. In ESL circles, however, the notion clashes with what is commonly taught, namely, that content words or loud function words create peaks right up to the primary stress. The fact is, they don't. And neither do ESL teachers I have listened to, including myself! Our natural inclination is to do in ESL classes what we do in ordinary speech. We use upstream destressing without noticing it. If our inclination is to sound natural, then why not use that inclination and explicitly help learners to follow it in the direction of more authentic discourse pronunciation?

ABOUT THE AUTHOR

Wayne Dickerson is Professor of Linguistics at the University of Illinois at Urbana-Champaign where he teaches pedagogical phonology in the MATESL program and ESL pronunciation for international students. He is certified as a Master Online Teacher and teaches the pedagogical phonology course online, as well. An applied phonologist, his principal interest is in the application of technical linguistic insights to practical teaching tasks, with an emphasis on English stress, rhythm, orthography, and materials development. His most recent books are *Stress in the Speech Stream: The Rhythm of Spoken English* (1989), published by University of Illinois Press, and (with co-author Laura D. Hahn) *Speechcraft: Discourse Pronunciation for Advanced Learners* (1999) and two workbooks, published by The University of Michigan Press.

REFERENCES

Beisbier, B. (1995). *Sounds Great! Intermediate pronunciation and speaking for learners of English, Book 2.* Boston, MA: Heinle & Heinle.

Bing, J. (1980). *Aspects of English prosody.* Bloomington, IN: Indiana University Linguistics Club.

Bolinger, D. (1986). *Intonation and its parts.* Stanford, CA: Stanford University Press.

Hahn, L. & Dickerson, W. (1999). *Speechcraft: Discourse pronunciation for advanced learners.* Ann Arbor, MI: The University of Michigan Press.

Gilbert, J. (2005). *Clear speech: Pronunciation and listening comprehension in North American English.* (3rd ed.) New York: Cambridge University Press.

Prator, C., & Robinett, B. (1985). *Manual of American English pronunciation*. (4th ed.) New York: Holt, Rinehart, and Winston.

Well, J. (2006). *English intonation: An introduction*. Cambridge, England: Cambridge University Press.

Wong, R. (1987). *Teaching pronunciation: Focus on English rhythm and intonation.* Englewood Cliffs, NJ: Prentice Hall Regents.

Baker, A. A. (2011). ESL teachers and pronunciation pedagogy: Exploring the development of teachers' cognitions and classroom practices. In. J. 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.

ESL TEACHERS AND PRONUNCIATION PEDAGOGY: EXPLORING THE DEVELOPMENT OF TEACHERS' COGNITIONS AND CLASSROOM PRACTICES

Amanda A. Baker, University of Wollongong

ABSTRACT

Over the past few decades, increasing more research has examined the cognitions (knowledge and beliefs) of second language (L2) teachers. Such research has provided insight into what constitutes teachers' beliefs and knowledge about teaching, how these cognitions have developed and how they are reflected in classroom practice (see Borg, 2006). Although numerous studies have been conducted into the curricular areas of grammar and, to a lesser extent, reading and writing, even fewer have examined teachers' cognitions into pronunciation instruction. The purpose of the present study, therefore, is to explore some of the dynamic relationships that exist between L2 teachers' cognitions and their actual pedagogical practices, but with a specific focus on how these cognitions have developed over time. In particular, the cognitions and practices - as they relate to the teaching of pronunciation - of five experienced ESL teachers are investigated. Overall, findings reveal that the amount of training teachers have received in pronunciation pedagogy strongly affects not only their knowledge of pronunciation and pronunciation pedagogy, but also their confidence in that area. Results further show that L2 learning experiences, teaching experience and collaborative work with colleagues can also influence teachers' practices and cognitions.

INTRODUCTION

One of the important areas of study in second language (L2) teacher education research is the study of *teacher cognition*. Exploring *teacher cognition* - defined as the relationship between the beliefs, knowledge and perceptions that teachers' have with respect to their teaching practices (Borg, 2003) - can provide insight into how teachers' beliefs and knowledge interact in the language classroom and influence their pedagogical behaviors, actions and activities. An essential element of this research, however, is the inclusion of observations of teachers' actual classroom practices and not merely teachers' self-reports of their practices (Borg, 2006). Although second language teacher cognition (L2TC) research has examined teachers' cognitions in relation to teaching many L2 skills areas, pronunciation is relatively underexplored.

L2TC research has focused mainly on the areas of grammar (e.g., Popko, 2005) and, to a lesser extent, reading (e.g., Johnson, 1992) and writing (e.g., Farrell, 2006). Many of these studies have explored connections between teachers' beliefs and actual classroom practices, finding that relationship between beliefs and practices may either correlate strongly (as in the case of Johnson, 1992) or demonstrate several differences between the two (as in the case of Farrell, 2006). Another area of primary interest in L2TC research is the investigation of different factors

that contribute to the development of teachers' cognitions. These factors have included teachers' previous experiences learning an L2 (Ellis, 2006; Farrell, 1999), teacher training (Borg, 1998; Popko, 2005), teaching experience (Gatbonton, 2008; Mattheoudakis, 2007) and knowledge sharing with colleagues (Sengupta & Xiao, 2002).

In comparison with L2 skills areas such as grammar and literacy mentioned above, pronunciation has received little attention. One of the few studies that has focused on teachers' knowledge and beliefs as related to teaching pronunciation is Baker (forthcoming); yet, this research does not include observations of the teachers' classroom practices. (See Baker & Murphy, forthcoming, for an expanded discussion of L2TC research and pronunciation pedagogy). This lack of research into pronunciation and L2TC is surprising considering the essential role that pronunciation plays in successful communication. Intelligible pronunciation is important not only in interactions involving native speakers, but between non-native speakers as well (Levis, 2005; Pickering, 2006).

In a similar vein, relatively few teacher education programs provide courses on how to teach L2 pronunciation. In fact, research has indicated that many L2 teachers have received little or no specific training in this area (Breitkreutz, Derwing & Rossiter, 2001; Derwing, 2010; Derwing & Munro, 2005; Murphy, 1997) and that teachers can be reluctant to teach pronunciation due to lack of training in pronunciation pedagogy and/or access to appropriate materials (Macdonald, 2002). Despite this apparent neglect, however, there is a demand among ESL learners for pronunciation instruction (Couper, 2003; Derwing & Rossiter, 2002). In fact, many learners hope to acquire native-like accents (Derwing & Munro, 2003; Kang, 2010; Scales, Wennerstrom, Richard, & Wu, 2006; Timmis, 2002) despite continued efforts to legitimize English as a Lingua Franca varieties (Jenkins, 2007) and other native models.

With the need for more L2TC research into pronunciation pedagogy (hereafter PrP) established, the current study focuses on one specific area of L2TC that warrants greater investigation, namely, the factors that contribute to the development of teachers' cognitions and pedagogical practices as related to pronunciation instruction.

RESEARCH QUESTIONS

This study provides a detailed description of some of the dynamic relationships that exist between experienced ESL teachers' cognitions and their observed practices when teaching pronunciation, focusing specifically on the development of these cognitions and practices. The research questions are:

1) What cognitions do experienced teachers have in relation to teaching pronunciation?

2) How have these cognitions developed?

3) How are these cognitions reflected in their actual practices?

RESEARCH DESIGN

Participants

Interviews and classroom observations were completed with five teachers in an Intensive English program (IEP) in a university in southeastern USA. For each participant, Table 1 provides: 1) years of teaching experience; 2) the level of oral communication (OC) course taught; 3) whether they have a degree related to TESOL; and 4) the location where that degree was obtained (Designated as Universities A, B and C, respectively). All names are pseudonyms.

Table 1: *Participant Backgrounds*

	Tanya	**Laura**	**Abby**	**Ginger**	**Vala**
Years Teaching Experience	7	6	6	14	7
Current OC Course	High Beginning	Low Intermediate	Intermediate	Intermediate	High Intermediate
MA degree – TESOL-related	Yes	Yes	Yes	Yes	Yes
Degree Location	A	A	A	B	C

Teacher Data

The teachers participated in three types of data collection procedures over one semester: three semi-structured (SS) interviews, five classroom observations, and two stimulated recall (SR) interviews. Figure 1 provides a general timeline for these procedures. The three SS interviews took place at different points, corresponding to the beginning, middle and end of the semester. The purpose of these interviews was to explore teachers' knowledge and beliefs about pronunciation and PrP. In the SR interviews, participants viewed pronunciation-related segments from the video-recorded observations and then recounted their cognitions at the time the event took place. The interviews and observed classes were transcribed.

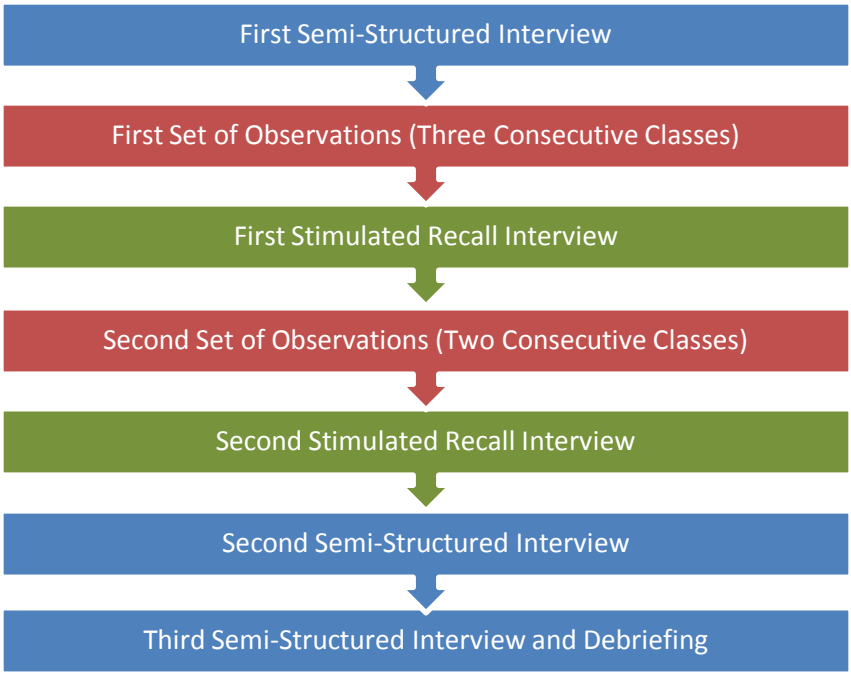


Figure 1: Research Timeline

FINDINGS AND DISCUSSION[1]

This section first examines how the teachers' knowledge of PrP has developed over time, and later explores this development in relation to the teachers' confidence in teaching pronunciation.

Development of Knowledge about Pronunciation Pedagogy

An integral component of trying to determine the source of teachers' knowledge of PrP required asking the teachers to describe their most prominent memories as they pertain to both learning an L2 and earning a TESOL degree. A part of this analysis included a brief look at their experience in teaching their respective OC courses. Table 2 provides a summary of these results.

Table 2: *Factors contributing to teachers' knowledge base of pronunciation pedagogy (PrP)*

	Tanya	**Laura**	**Abby**	**Ginger**	**Vala**
Language learning experience	French; Spanish	Spanish, Latin	Bilingual (English & Port); Norwegian	French, Spanish, Japanese, Turkish	Spanish
--Most memorable component related to pronunciation	Repetition: sounds/words	Spanish phonology course (sound formation)	Repetition: Words	Repetition: Words and Phrases	Repetition: Words and Sounds
MA course that focused the most on PrP	Pronunciation	Pronunciation	Pronunciation	Listening/Speaking	Methods
--Textbook	Celce-Murcia et al. (1996); Grant (2001)	Celce-Murcia et al. (1996)	Celce-Murcia et al. (1996)	Celce-Murcia et al. (1996)	possibly Ur (1991)
--Most memorable component	Phonetic symbols/ individual sounds	Classroom practices – design activities for students	Techniques: Word stress & Rhythm	Analyzing learner language (tutoring)	Listen to Group micro-teaching (Chinese)
Instructor	W	X	X	Y	Z
Current OC course	High Beginning	Low Intermediate	Intermediate	Intermediate	High Intermediate
Times taught	1	1	6-7	3	4
Features taught in this course	Syllables, word endings, consonants, vowels, word stress, rhythm, intonation, connected speech	Vowels, consonants, syllables, word endings, word stress	Syllables, rhythm, vowels, word stress	Syllables, rhythm, vowels, word stress	Syllables, word stress, word endings

As illustrated in Table 2, the five participants have experience learning at least one L2 as part of their secondary and tertiary educations. Of the five participants, only one was bilingual, Abby, who grew up learning both English and Portuguese. When asked to describe the most memorable moments learning pronunciation in an L2, the instructors' descriptions were, for the most part, very similar. In every case, the repetition of sounds, words and/or phrases represented the essence of learning L2 pronunciation for these teachers. The only participant who differed was Laura who, as part of her bachelor's degree in Spanish, took a course in Spanish phonology, where she learned about the articulation of sounds.

Specific training in PrP differed among the instructors. Altogether, they received MA degrees from three different universities. The graduate curriculum at institution A included a course that was entirely devoted to PrP. Tanya took the course with instructor W while Laura and Abby took it with instructor X. At institution B, Ginger took a course on teaching listening and speaking which included a component on pronunciation instruction. At both of these two institutions, three graduate course instructors, all of whom are well-known specialists in pronunciation instruction, used Celce-Murcia, Brinton, and Goodwin's (1996) book in their respective courses. Instructor X also used Grant's *Well said* (2001). Vala, unlike the other four instructors, received little or no education in PrP. The closest training she received was the small amount that was covered in a course on TESOL methodology.

When asked to describe the most memorable part of their PrP course, the teachers' responses varied considerably, ranging from learning about the articulation of individual segmentals to learning specific techniques for teaching prosodic elements to practice working with an English language learner. Tanya responded that being introduced to phonetic symbols and descriptions of individual sounds had the largest impact on her. For Laura, she attached more importance to classroom practices in general and how to design specific pronunciation-oriented activities for students. In Abby's case, she best recalled specific techniques for teaching word stress and rhythm. Ginger, who took a course on teaching listening/speaking, was only able to remember one, albeit major, PrP component of that course. For her, she remembered a tutoring project where she had to analyze the speech of her student, identify three features of English pronunciation that he had the most difficulty with (done in consultation with course instructor Y), and work with the student to improve his pronunciation in those three areas. Finally, Vala's situation differed the most from the other four instructors. From what she was able to recall, only the Methods had a focus on pronunciation instruction, and that focus solely involved student microteaching sessions in which one of the groups taught a mini-lesson on Chinese pronunciation.

Another major factor that is frequently cited in the literature as having a significant influence on teachers' knowledge about language teaching, and that also has a major influence on the teachers in the current study, is their teaching experience (Gatbonton, 2008; Mattheoudakis, 2007). Each instructor has experience teaching the specific OC course investigated in the present study, thus through teaching the required content of this course, they have either learned about or enhanced their previous understanding of different elements of English pronunciation. Both Tanya and Laura have taught their respective OC courses once, Abby has taught her course six to seven times; Ginger has taught her course three times; and Vala has taught her course four times. Table 1 lists the features of pronunciation taught in each course. (These same features are also required

elements to teach in the IEP curriculum). As will be shown later, this experience has played a role in both their knowledge of and confidence in teaching pronunciation.

Strongest Influence on Current Practice

Another component of this study focused on one of the questions asked in the interviews: *What has had the greatest influence on how you teach pronunciation?* Both the teachers' responses and their observed classroom practices are explored in addressing this question.

Tanya

According to Tanya, the factors that have had the strongest influence on how she teaches pronunciation today is her graduate education, learning through "trial and error" and "having a good textbook to follow". This latter factor can be directly linked to her graduate education. As part of her PrP course and during her MA practicum placement, she used Linda Grant's (2001) *Well Said*; thus, she received a "double dose" of sorts of the material covered in this one book.

Her current classroom practice in many ways reflects this education. The observations revealed that Tanya regularly uses Linda Grant's (2007) *Well Said Intro* with her high-beginning course. Many of the explanations that Tanya gives about rules or guidelines related to pronunciation come from this text as do the majority of activities that she uses in the classroom. She explained that she "really like[s]" the activities in this book and that this book is "the basic standard". Using the textbook, however, does not represent the entirety of her lessons. She also occasionally supplements materials or activities that she acquired in her graduate course in PrP.

Laura

As with Tanya, Laura also identified her graduate education as having the greatest impact on how she teaches pronunciation in her low-intermediate OC course. She noted that "I think if I wouldn't have had that [PrP] course…I think I would feel very uncomfortable teaching pronunciation… [It's] probably had the most effect as far as me actually teaching pronunciation and how I teach it."

Observations of her lessons corroborated Laura's belief that her graduate education played a key role in how she taught pronunciation to her student. During these observations, Laura used several techniques that she mentioned having learned as part of her pronunciation pedagogy course. One of these included stretching a rubber band when saying the stressed part of a word (to demonstrate word stress). She also supplemented her lessons with exercises from a pronunciation textbook. She also supplemented her lessons with exercises from Gilbert's *Clear Speech* (2005).

Abby

Similarly, Abby also considered her graduate education as having the most influence on her teaching of English pronunciation. She stated, "I think the biggest influence is [Instructor

Y]…and that is where I get most of …my ideas, knowledge." She explained that learning that there are rules for stress in English was "quite eye-opening."

The observed classes reflected this previous training. Many of the techniques that she reported learning in the MA course surfaced in the observed classes. These techniques included, but were not limited to, using both Acton's "syllablettes" (see Acton, 1998) and rubber bands (see Gilbert, 2008) for teaching word stress.

Ginger

Unlike the previous three participants, Ginger considered her experiences of learning multiple foreign languages as having the strongest influence on her pronunciation teaching, although collaboration with a colleague also played a prominent role in her teaching. In particular, she discussed the importance of not singling students out in class, especially when working on pronunciation. Avoiding putting students in situations where they might feel embarrassed was a principle that she highlighted as having priority in her classes. Furthermore, she reported remembering only a limited amount from her graduate coursework. As mentioned earlier, her main memory of that course centers on the tutoring project she did. She explained that the focus of the MA course was on listening and speaking, resulting in less time being devoted to pronunciation. In addition, she furthered identified collaboration with Abby as having an impact on how she teachers. Referring to the time when she taught the intermediate OC course for the first time, Ginger said, "I went out of my way to bug Abby…Abby, what's this? What's that? Why? Why do you use the kazoo? …We were very collaborative …"

Ginger's observed practice mirrored her beliefs, but not to the same extent as she reported in the interviews. Observations showed the impact that her earlier language learning experiences had on her teaching: she rarely singled out students in front of the class to give them feedback on their pronunciation. However, while Ginger emphasized the impact of her experiences as a L2 learner when asked what had the greatest influence on her teaching, what seemed to have an even greater impact was her collaboration with Abby. As I observed both Abby and Ginger teach pronunciation, many of the activities I saw in Abby's classes, one's that Abby had pinpointed as activities or techniques she had learned in her graduate course on pronunciation pedagogy, also appeared in Ginger's lessons. In fact, several of these techniques Ginger mentioned as having learned from Abby. Thus, in many ways, the graduate education received by Abby also benefited Ginger's teaching of L2 pronunciation. Ginger's collaboration with Abby, has had considerable influence on how Ginger teaches pronunciation to her students.

Vala

The development of Vala's knowledge of PrP differed from the other instructors. When asked what has had the most influence on her teaching of pronunciation, Vala responded with "I don't know. That's hard to say. I think all of that experience of "just winging it" taught me a lot…all of those crazy experiences I've had of playing teacher without any guidance were pretty influential." In this quote, she is referring, at least in part, to her experience as an MA student trying to complete an unsupervised practicum that she did in a volunteer ESL program. Without

any training in how to teach pronunciation and having no guidance from a practicum supervisor, she felt as though she had to learn everything on her own.

During the observations, Vala showed a strong dependence on the course textbook when teaching pronunciation, a book that was not devoted to pronunciation, but rather to OC skills in general. For the most part, she adhered to the limited number of pronunciation activities in that book; however, she made the activities more interactive. She had the students work in groups and she encouraged them to use each other as models and teachers of pronunciation, explaining to them that she "will not always be [their] teacher".

All instructors

The stories described above show that graduate education, collaboration with colleagues, textbooks, and teaching experience have had an impact on the five teachers to varying degrees. For Tanya, Laura and Abby, the interviews and observations revealed that graduate education in many ways seemed to have the strongest influence on how they teach pronunciation, whereas for Ginger, who had less training, and for Vala, who had no training, formal education seemed to play a minimal role, if at all. These pronunciation-related results are similar to the grammar-related findings of Borg (1998) in that Borg found that teacher training seemed to override one teacher's prior beliefs about language teaching. In the present study, for teachers with limited or no teacher training in PrP, the textbook and/or collaboration with a colleague appeared to have the greatest impact on their teaching of pronunciation. Ginger learned more about teaching pronunciation from working with a colleague. This notion of knowledge sharing with colleagues has further been found to aid teachers to learn from their teaching experiences in other research (Sengupta & Xiao, 2002). Finally, for Vala she appeared to learn how to teach pronunciation mainly from ESL textbooks and from teaching experience.

In addition to the presence or absence of graduate training in PrP, classroom observations and/or interviews revealed that prior L2 learning experience, namely that of repetition drills, also had a role in the teachers' classroom practices. During the interviews, each of the five teachers identified repetition work as the most memorable activity, or the only activity, that they recalled of learning pronunciation in an L2. Similarly, the role of repetition work was apparent in each of observed OC courses and/or discussed by the teachers during the interviews. Repetition work was evident in many of the observed classes taught by Tanya, Laura and Vala; however, for Abby and Ginger, links to prior L2 learning experience could only be based on the teachers' self-reported use of repetition drills in their OC courses. Nevertheless, observations revealed that Abby's and Ginger's students experienced a more delayed type of repetition practice. Instead of repeating immediately after their teacher, the students reproduced the teacher's utterances (whether a sound, word, phrase or sentence) several seconds or even up to a minute after hearing the utterance. The teachers would ask the students to reproduce an utterance typically after the teacher either pointed to a written word/sentence or a picture) or used a gesture or other kinesthetic movement (e.g., clapping to a measured beat) to indicate that students were expected to reproduce the earlier teacher's utterance.

Confidence and Pronunciation Instruction

Strongly connected to teachers' knowledge about teaching pronunciation is their confidence in that knowledge. I asked each of the teachers: *How confident are you in teaching pronunciation?* Their responses are provided in Table 3:

Table 3: *Teachers' Confidence in Teaching Pronunciation*

Participant	Confidence in teaching pronunciation
Tanya:	"Some days I feel more confident than others. Probably has more to do with my personal feelings versus my ability to teach pronunciation...But generally, especially if I've taught the class before, I feel pretty certain I can answer any questions that come up or figure out the answer, make up an answer to a question."
Laura:	"I think I am confident, fairly confident teaching it, maybe 7 out of 10, but sometimes I worry what are the students getting out of this. I am not sure I guess that my instruction is really making a difference down the line."
Abby:	Teaching level 3 yes. And actually I think if I were to go down now, if I were to go down to level 2 or level 1, after this project, you know, we were really working on segmentals [in reference to her research project with Ginger] yeah. Yeah, now I feel like I could handle level one and level two.
Ginger:	"I'm much more confident now. If you asked me spring of 2008, I would have been like I'm just a disaster…But now that I've done it a few times and I know the subject matter of the course really well, I guess I should say in the context of my course, I feel really confident."
Vala:	"I don't have enough experience. I feel that I could be better at it. I would be better at it if we had a pronunciation class, because I'd get the experience. But then, the classes that we teach don't really incorporate that much, so outside of the stuff that gets covered in our classes and like extemporaneous stuff that happens, I don't feel so confident with it. I don't feel that I was exposed enough to it, to how to teach it different ways and have to get them to practice it different ways. Beyond what I do, not so confident. New territory, a little scary!"

The thoughts expressed by each of the participants indicate that graduate training in PrP and experience teaching a specific course may have had a direct impact on their confidence in teaching pronunciation. The four teachers with graduate training in PrP all expressed having confidence in teaching pronunciation at least in the OC course in which they have experience teaching. For Vala, however, the story is different. Even though she had taught the course four times in the past, she remains insecure in her teaching of pronunciation. Beyond the few techniques that she has learned for teaching pronunciation features in her course, she keenly feels the gaps in her knowledge base as a result of her lack of training. Nevertheless, even the teachers with training in PrP are noticeably uncertain in their ability to teach new features of pronunciation not covered in their courses. However, unlike the findings of MacDonald (2002)

that showed that inadequately teachers were reluctant to teach pronunciation, neither the classroom observations nor the interviews indicated that the five teachers were even partially reluctant to teach pronunciation. In fact, Vala expressed a desire to teach a pronunciation course in order to expand her knowledge base of PrP.

CONCLUSION

Although the findings from this study are based on the cognitions and practices of only five teachers, it seems clear that TESOL programs comprising at least one course dedicated to PrP may have a significant impact on both teachers' knowledge of PrP and confidence in teaching pronunciation. Furthermore, unlike grammar and literacy skills, pronunciation appears to be rarely or only partially taught in L2 learning experiences of teachers; thus, many teachers may have neither adequate knowledge about English pronunciation nor sufficient knowledge of how to teach it. Overall, these results indicate that increased attention needs to be devoted to PrP in TESOL programs.

ABOUT THE AUTHOR

Amanda Baker is a Lecturer in the Faculty of Education, TESOL program at the University of Wollongong, NSW, Australia. She is also an adjunct instructor in the MA in TESOL program at Trinity Western University (British Columbia, Canada) where she teaches a course in pronunciation pedagogy. Her research interests include second language teacher education, classroom research and pronunciation pedagogy.

NOTE

[1] The results reported here are based on partial findings from the author's dissertation project on *Pronunciation pedagogy: Second language teacher cognition and practice*. The larger project also includes an examination of students' beliefs in comparison with the beliefs of the teachers.

REFERENCES

Acton, W. (1998). Directed motion in second language pronunciation instruction. *Journal of Communication and International Studies, 5*(1), 43-52.

Baker, A. A. (forthcoming). Discourse prosody and teachers' stated beliefs and practices. *TESOL Journal*.

Baker, A. A. & Murphy, J. (forthcoming). Knowledge base of pronunciation teaching: Staking out the territory. *TESL Canada Journal*.

Borg, S. (1998). Teachers' pedagogical systems and grammar teaching: A qualitative study. *TESOL Quarterly*, *32*(1), 9-38.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe and do. *Language Teaching, 36*(2), 81-109.

Borg, S. (2006). *Teacher cognition and language education: Research and practice*. London: Continuum.

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

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

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect, 18*(3), 53-70.

Derwing, T. M. (2010). Utopian goals for pronunciation teaching. In J. M. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (pp. 24-37). Ames, IA: Iowa State University.

Derwing, T. M., & Munro, M. (2003). What do ESL students say about their accents? . *The Canadian Modern Language Review, 59*(4), 547-566.

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., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System, 30*, 155-166.

Ellis, E. M. (2006). Language learning experience as a contributor to ESOL teacher cognition. *TESL-EJ, 10*(1), 1-20.

Farrell, T. S. C. (1999). The reflective assignment: Unlocking pre-service teachers' beliefs on grammar teaching. *RELC Journal, 30*(2), 1-17.

Farrell, T. S. C. (2006). Reflective practice in action: A case study of a writing teacher's reflections on practice. *TESL Canada Journal, 23*(2), 77-90.

Gatbonton, E. (2008). Looking beyond teachers' classroom behaviour: Novice and experienced ESL teachers' pedagogical knowledge. *Language Teaching Research, 12*(2), 161-182.

Gilbert, J. B. (2005). *Clear speech: Pronunciation and listening comprehension in North American English. Student's book* (3rd ed.). New York: Cambridge University Press

Gilbert, J. B. (2008). *Teaching pronunciation: Using the prosody pyramid*. Retrieved from http://www.cambridge.org/other_files/downloads/esl/booklets/Gilbert-Teaching-Pronunciation.pdf

Grant, L. (2001). *Well said: Pronunciation for clear communication* (2nd ed.). Boston, MA: Heinle & Heinle.

Grant, L. (2007). *Well said intro: Pronunciation for clear communication*. Boston, MA: Thompson Heinle.

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

Johnson, K. E. (1992). The relationship between teachers' beliefs and practices during literacy instruction for non-native speakers of English. *Journal of Reading Behavior, 24*, 83-108.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3), 369-378.

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.

Macdonald, S. (2002). Pronunciation views and practices of reluctant teachers. *Prospect, 17*(3), 3-18.

Mattheoudakis, M. (2007). Tracking changes in pre-service EFL teacher beliefs in Greece: A longitudinal study. *Teaching and teacher education, 23*(8), 1272-1288.

Murphy, J. (1997). Phonology courses offered by MATESOL programs in the US. *TESOL Quarterly, 31*(4), 741-764.

Pickering, L. (2006). Current research on intelligibility in English as a lingua franca. *Annual Review of Applied Linguistics, 26*(1), 219-233.

Popko, J. (2005). How MA-TESOL students use knowledge about language in teaching ESL classes. In N. Bartels (Ed.), *Applied linguistics and language teacher education* (pp. 387-404). New York: Springer.

Scales, J., Wennerstrom, A., Richard, D., & Wu, S. H. (2006). Language learners' perceptions of accent. *TESOL Quarterly, 40*(4), 715-738.

Sengupta, S., & Xiao, M. K. (2002). The contextual reshaping of beliefs about L2 writing: Three teachers' practical process of theory construction [Electronic Version]. *TESL-EJ, 5*. Retrieved May 17, 2008 from http://cwp60.berkeley.edu/tesl-ej/ej21/a1.html.

Timmis, I. (2002). Native-speaker norms and international English: A classroom view. *ELT Journal, 56*(3), 240-249.

Ur, P. (1991). *A course in language teaching: Practice of theory*. Cambridge: Cambridge University Press.

CONTACT INFORMATION

Amanda A. Baker
Lecturer in the Faculty of Education, TESOL program
University of Wollongong
Wollongong NSW 2522 Australia
aabaker20@hotmail.com

Reed, M. & Michaud, C. (2011). An integrated approach to pronunciation: Listening comprehension and intelligibility in theory and practice. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (pp. 95-104), Ames, IA: Iowa State University.

AN INTEGRATED APPROACH TO PRONUNCIATION: LISTENING COMPREHENSION AND INTELLIGIBILITY IN THEORY AND PRACTICE

Marnie Reed, Boston University

Christina Michaud, Boston University

This paper introduces a theoretical basis for understanding the relationship between speaking and listening as an auditory feedback loop, in which speakers use their own mental model of a sound as input. We argue that speaking and listening are related: production facilitates perception, promoting more intelligible spontaneous speech and enhanced listening comprehension. Since learners use their own speech production as a filter, we suggest a model based on convergent production. The model is compared to Flege's Speech-Learning Model and also applied to the practical problem of pronunciation instruction.

We identify four major components of pronunciation instruction: connected speech features, suprasegmental features, inflectional morphology, and segmentals. Within this integrated model of pronunciation, the route to successful listening comprehension is through auditory feedback wherein the learner's own increasingly target-like speech production facilitates and reinforces perception. We introduce specifics for promoting learners' convergent output: spoken models, visual aides, and oral or written descriptions. A teacher-student partnership in which teachers offer a principled approach to corrective feedback in the form of production prompts is advocated. In our model, pronunciation instruction accompanies and reinforces core language instruction, and integrated pronunciation instruction is viewed as a highly focused, metacognitive approach to the entire language classroom.

INTRODUCTION

Research findings and anecdotal observations alike indicate that adult second language learners often have intractable pronunciation problems that persist in the face of explicit instruction or correction. Mere target language input is insufficient to create changes (Flege & Hillenbrand, 1984; Flege, 1993; Strange, 1995), or learners would never need instruction. The long-established failure of adult learners to incorporate syntactic structures to which they are exposed (Pavesi, 1986; Long, Inagaki, & Ortega, 1998) has a parallel in pronunciation. Learners frequently seem not to "hear" the target pronunciation even when it is modeled by their teachers, instead continuing with their original, incorrect pronunciation.

One way of making sense of this problem is to conceive of the relationship between a learner's speaking (production) and listening (perception) as an auditory feedback loop, in which speakers use their own output—their own mental model of a sound—as input for their production. The role of auditory feedback, attested in research on speakers with normal hearing and profound hearing loss (Perkell, Matthies, Lane, Guenther, Wilhelms-Tricarico, Wozniak, & Guiod, 1997; Perkell, Guenther, Lane, Matthies, Perrier, Vick, Wilhelms-Tricarico, & Zandipour, 2000) also

helps account for the persistence of both spoken error and a lack of aural discrimination between the target sound and the error.

The question then arises of how to break this feedback loop; naturally enough, it can be broken when a speaker actually forms the target sound in his or her mouth and then combines this new and different motor-memory with a new acoustic image of the sound. In practical terms, we tell learners that when they finally "get it right"—produce a target sound—it will "feel wrong" to them, but they should nonetheless perform a mental freeze-frame at that moment.

For our purposes, it doesn't matter at all how to get speakers to produce the sound: teachers may use any means at their disposal, including aural models, mouth diagrams and mirrors, metalinguistic feedback, and pronunciation aids, such as rubber bands, Chinese yo-yos, and kazoos (Gilbert, 2005; Grant, 2004; Hewings, 2004). We believe that aural models will be less helpful, by and large, than other means, but teachers should on no account rule out any particular way of getting learners to produce the target sound (Couper, 2006; Gilbert, 2010). We agree with Flege (1993, 1995) that, when applicable, the contrasting nature of L1 and L2 features may help learners in their acquisition, but we argue that teachers must simply be more creative and proactive to help learners acquire pronunciation features that may not be as saliently different from their L1 analogs.

Often speakers *can,* in fact, produce the target sound (Fraser, 1999, 2000), but in limited phonological environments. Word-initial /w/ in "want or "what" is not problematic for Japanese speakers; however, since /ʊ/ does not occur in Japanese, word-initial /w/ in "would" or "woman" is difficult. By helping the speaker to isolate the target sound in a licensed consonant-vowel combination, we can start to help the speaker to form a new mental model in expanded CV syllable contexts.

Once teachers can get a learner to produce the target sound, in any context, and on the road to creating a new mental model of it, both the learner's speaking and listening are expected improve. We believe that in a model of convergent production, production precedes and facilitates perception, promoting more intelligible spontaneous speech and enhanced listening comprehension as the speaker's production begins to converge on the target: as we tell learners, speaking helps listening. We posit that our production-centered approach to intelligibility is actually consistent with Flege's Speech-Learning Model (1993, 1995). Flege's SLM focuses on the mental and metacognitive work of perceiving a new target sound; our model does not contradict the SLM, but instead offers a specific and practical pedagogy to allow learners to create acoustic images of new sounds.

AN INTEGRATED MODEL

With what we see as this key relationship between speaking and listening in mind, therefore, we argue for an expanded model of pronunciation which includes both speaking and listening and addresses both suprasegmantals and segmentals. Figure 1 represents our Integrated Model of Pronunciation:

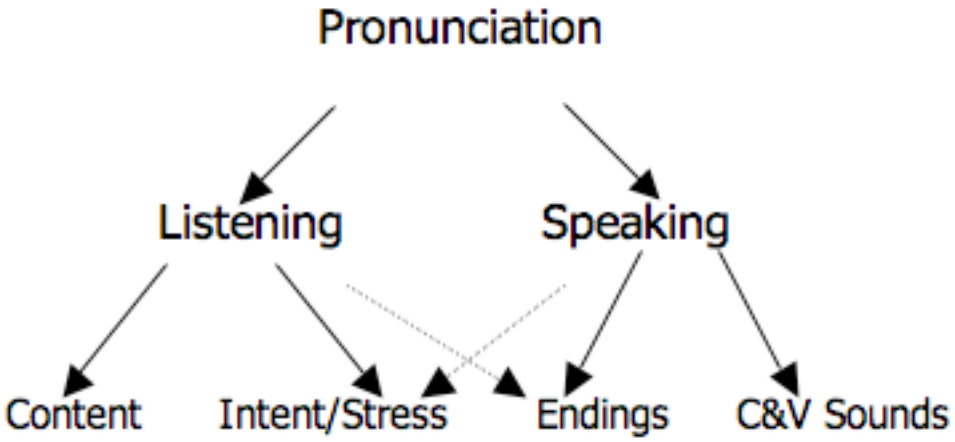


Figure 1. Integrated model of pronunciation

We identify four major components of pronunciation. Working from left to right in the figure above, we identify the first strand as listening for *content*, because learners who struggle with this strand of pronunciation tend to miss or misunderstand the actual content of what is being said. This strand contains the connected speech features—linked, reduced, deleted, altered, and contracted sounds. Though learners often think that the speed of English accounts for their listening problems, in fact these connected speech features are contributing sources. One of these features, reductions, has been frequently cited as a major contributing factor to misinterpretation of the homophones "to" and "two" in civil aviation (Waldock, 1994), resulting in a 1989 crash when an air traffic controller cleared an aircraft to descend to "two four zero zero" but the pilot interpreted the number "two" as the preposition "to" and descended accordingly (Cushing, 1995a, 1995b).

Next, we identify the strand of *intent*—the suprasegmental features of English (syllables, stress, intonation, rhythm, and timing) that convey speaker meaning. Learners who struggle with this aspect of pronunciation may understand every word a speaker says but may still miss the meaning; conversely, they may be unable to get their own content or intent across. This is why we show a dotted arrow coming down to this strand from speaking, as well; we argue that suprasegmentals are essential both for speaking and listening (Hahn, 2004). In a study currently being undertaken, advanced-level learners [Level D in an A-E level intensive English program, consisting of 14 students representing five native language groups: Arabic (n = 4), Korean (n = 4), Mandarin Chinese (n = 3), Spanish (n = 2), and French (n = 1)] were questioned prior to suprasegmental instruction on their opinions regarding the role of one component of suprasegmentals: intonation. Using a digital response system (clickers), students voted for their choice of response to the following question:

> Which is more important in a sentence?
>
> a. the words in the sentence
>
> b. the intonation in the sentence

For this question, 73% (*n* =11) of the respondents voted for choice a: words, while 27% (*n* =3) of the respondents said that intonation was more important. Following completion of the first lesson on intonation and contrastive stress in their pronunciation text (Grant, 2009), students were asked their interpretation of a sentence with non-standard sentence-level stress. Using the digital response system, learners responded to a yes or no question after hearing an audio prompt:

Audio Prompt: *The teacher* didn't grade your papers.

Question: Were the papers graded?

Yes

No

The votes split evenly (*n* = 7 'yes' votes, 7 'no' votes), may reflect some effects of instruction and generating much discussion. When the correct response, (a), was displayed, a number of learners asked to hear the sentence again. Several could be seen and heard repeating the sentence under their breath, particularly the verb phrase "didn't grade." These respondents demonstrated grasp of the lexical, morphological, and syntactic content of the utterance, but missed the intended meaning, suggesting that learners are unaware of the importance of—in this particular case—contrastive stress and emphatic intonation, and their role in altering the meaning of an utterance. Attention to this suprasegmental strand is therefore essential for learner success in listening as well as in speaking

Returning to Figure 1, the third strand from the left is *endings*—in other words, the sounds of regular verb and noun inflectional morphology, which cause learners serious and intractable problems when they omit or mispronounce these endings. The fact that learners fail to pronounce these endings on regular nouns and verbs when reading aloud convinces us that the problem is not always a local grammar error, but is based in their internalization of the sound system. Teachers, based on learners' ability to state the grammar rules and produce these endings in controlled drills and tests, appear to succumb to what Gass and Varonis (1985) refer to as accommodation, or not noticing the errors in learners' spontaneous speech. Whether teachers do in fact unconsciously accommodate—i.e., no longer notice learners' errors in this area—or instead make a conscious decision to avoid addressing a seemingly intractable problem, these errors are nevertheless serious ones. Studies have shown that seemingly small errors with these endings nevertheless stigmatize learners and are responsible for communication breakdowns (Jiang, 2007; Lardiere, 1998; Long, 2003; Major, 1995, 1998). The misinterpretation of "He looked it up" as "He looked up" by a speaker who pronounces "look" plus past tense as two syllables attests to the role of the acoustic image and the interrelation of production and perception. For these reasons, we believe that this strand, too, is key for both learners' speaking and listening skills, as represented in Figure 1.

Finally, our fourth strand of pronunciation is that of individual *consonant and vowel sounds* (segmentals) which learners typically report to be their biggest pronunciation errors and main source of communication breakdowns (Derwing & Rossiter, 2002). Our experience in a pronunciation elective class with a Japanese neuroscientist who did not differentiate between liquids /l/ and /r/ or fricative /ʒ/ and affricate /dʒ/, rendering "brain regions" indistinguishable

from "brain lesions," attests to the need to address segmentals. Nevertheless, we weight this strand as relatively less important for intelligibility than the two strands toward the middle of the figure. Thus our model advocates a combined focus on segmental and suprasegmental features and an integrated approach to instruction.

THE MODEL IN CLASSROOM PRACTICE

We believe that this integrated model of pronunciation is particularly helpful to learners for making sense of their errors: teachers are able to break the large topic of pronunciation down into meaningful, yet discrete chunks, and then build it back up as learners master various patterns and sounds. Teacher tools that can help learners have success with these different parts of pronunciation abound: we suggest a teacher-learner partnership, where teachers guide learners through both the production of these different pronunciation topics and the metacognition necessary for "noticing" these patterns (Schmidt, 1990) and building up new acoustic images of the patterns (Neufeld, 1977).

Three sample entries in a pronunciation logbook, along with directions written for learners (Figure 2), show one way that metalinguistic feedback can help move a learner toward more target-like production—in this case, of the problem of initial consonant clusters, a syllable structure problem we classify in the second strand (suprasegmentals, or intent/rhythm) above:

Use a log to help you enter mistakes, label them by kind, and review a list of *your* biggest problems. If you keep making the same kinds of errors, use your log to help you correct them.

Everyone has different pronunciation problems. Look at the examples on the chart below to see the different difficulties three students had with the same word. To improve, students need to know what their individual mistakes are.

Word:	How should I say it?	How did I say it?	What was my mistake?	Other examples:
Student 1: speech	speech (1 syllable)	su-peech (2 syllables)	separating the first two consonants	su-trong/strong
Student 2: speech	speech (1 syllable)	es-peech (2 syllables)	adding a vowel sound at the front	es-port/sport
Student 3: speech	speech (1 syllable)	speech-ee (2 syllables)	adding a vowel sound at the end	each-ee/each

Figure 2. Sample pronunciation logbook

In Figure 3, we offer a checklist for learners who "know" the rule for the pronunciation of the regular past-tense verb ending, but nevertheless omit it entirely or incorrectly pronounce it as two syllables in verbs like *looked* and *used*. These errors, of course, relate to the third strand (verb and noun endings, or morphosyntax) in our integrated model of pronunciation:

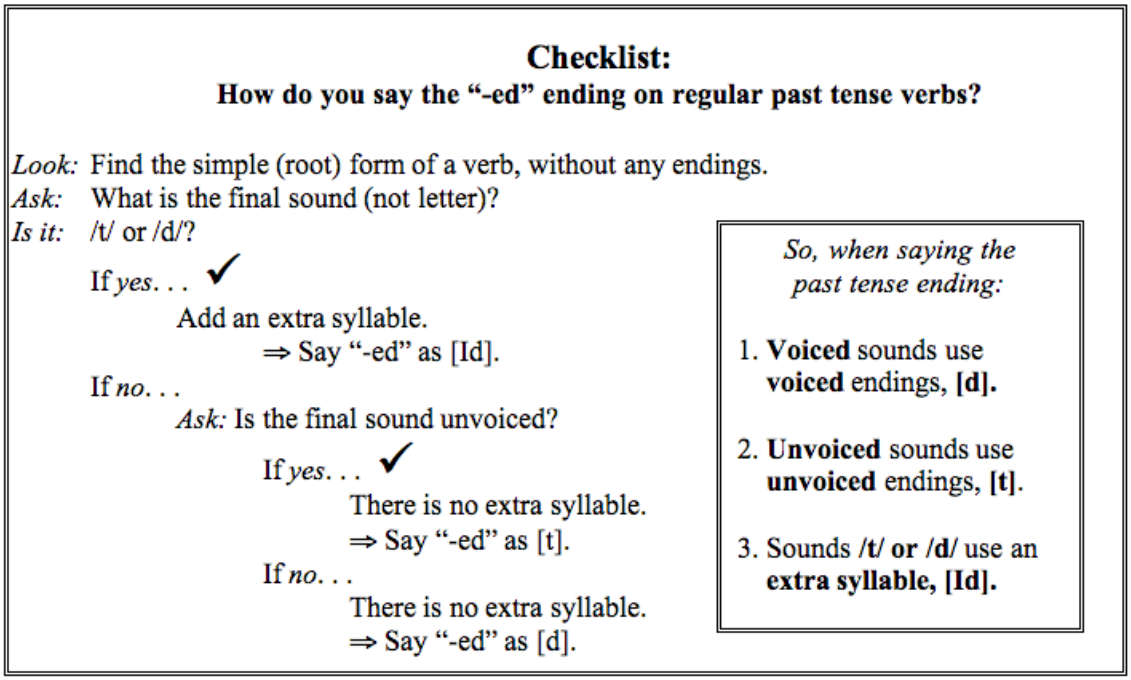


Figure 3. Checklist for pronouncing regular past tense endings

Teacher intervention is key to helping learners move from isolated production of a new pronunciation sound or pattern to full production of it in spontaneous speech, and in turn to full perception of it in other speakers' connected speech. Such intervention can be seen as the real work of the pronunciation classroom—it is what teachers are there to do, after all, and it is why we don't advise learners just to improve their pronunciation by merely reading a pronunciation book (even a good one). The model describes the intermediate stage of spoken proficiency when learners are able, sometimes, to produce a target sound or pattern on demand, but have not yet integrated it fully into their spontaneous speech as the stage of prompted production: through instruction in the sound or pattern, and careful prompting (again, the actual means of getting the learner to produce this sound may vary widely), learners will achieve their first success at producing the sound in a target-like fashion. Learners then move through a stage when teachers' corrective feedback is necessary to help them solidify the pattern in their own mouths and minds, before they fully form a new mental model of the sound or pattern. To better help learners, we suggest that the language of instruction match the language of correction; in other words, whatever metalinguistic feedback teachers wish to offer ("Say it as one syllable," "Don't use a voiced sound there," "Make the stressed syllable louder, longer, clearer, and higher," etc.) in the classroom, to individual learners, should be the same terms as those in which they teach the concept or pattern to begin with. That language should also be the same that teachers elicit from learners in the form of learner "tell-backs." Here we borrow the term from the literature on reading instruction (Vanderwood, 2007), where tell-backs are defined as classroom language that articulates, superficially for the teacher's benefit, but more significantly for the learner's own benefit, what the learner now knows about this concept (e.g., "I said it as two syllables, but I need to only say one; I need to omit the final syllable after the sibilant.") and what will help the

learner go about forming a new mental model. Over time, as learners demonstrate internalization and metacognitive awareness, external corrective feedback can take the form of increasingly precise, minimal, and unobtrusive 'production prompts' which allow the learner to self-correct and return to finish their thought. Figure 4 shows the path to intelligibility, and the connections among the different terms we have been using in this model:

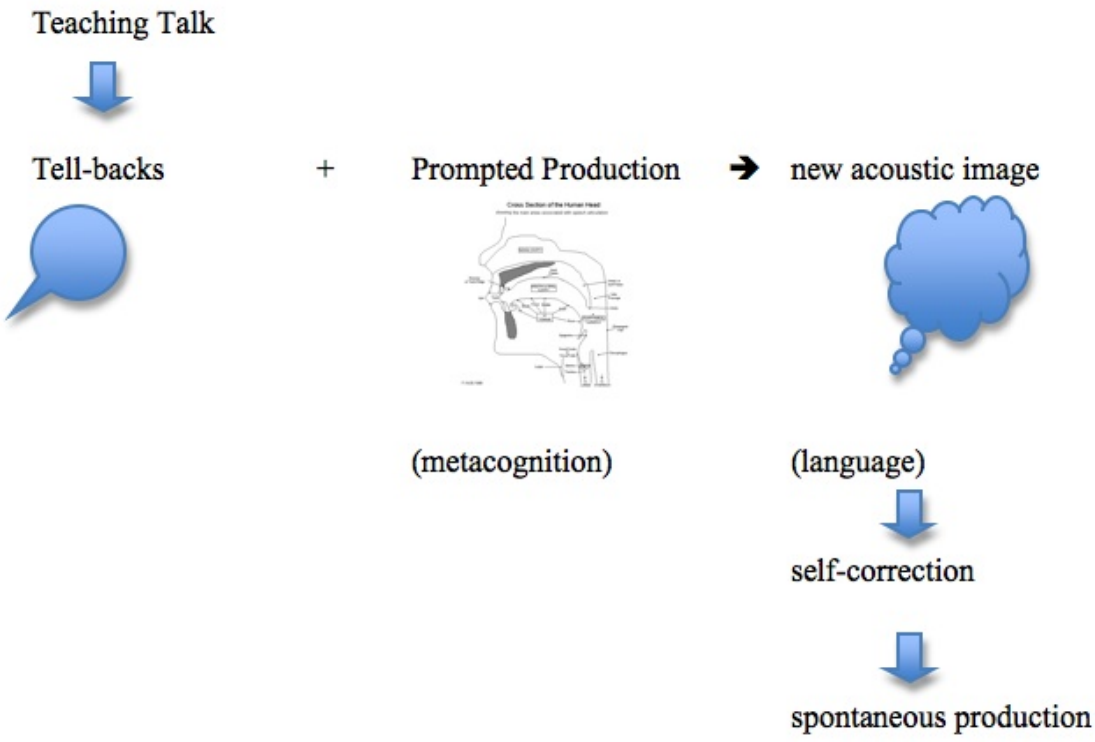


Figure 4. The path to intelligibility

In our model, pronunciation instruction accompanies and reinforces core language instruction. While Flege's SLM (1993, 1995) focuses primarily on segmentals, our model considers the acquisition of all aspects of pronunciation in a more integrated way; Flege's emphasis on the role of time, the nature of instruction, and the learner's internal representations of both L1 and L2 are also key to our understanding of the pronunciation acquisition process. Ultimately we view integrated pronunciation instruction as a highly focused, metacognitive approach to the entire language classroom, and a key step along the way to learners' increasing intelligibility.

ABOUT THE AUTHORS

Marnie Reed is an Associate Professor in the graduate TESOL Program and the Program in Applied Linguistics at Boston University. The focus of her research has been on second language acquisition, with a focus on the relationship between speech production and perception. She has

co-authored the books *Sound Concepts: An Integrated Pronunciation Course* (2005, McGraw-Hill) and its accompanying *Teacher's Manual* (2006), and *Goal-Driven Lesson Planning for Teaching English to Speakers of Other Languages* (2010, University of Michigan Press), which incorporates her work on pronunciation instruction.

Christina Michaud is a Lecturer in the College of Arts and Sciences Writing Program at Boston University and also a doctoral candidate in Language, Literacy, and Cultural Studies in the School of Education at BU. She teaches ESL composition as well as general first-year composition and research writing. She has co-authored two books and numerous conference presentations with Marnie Reed.

REFERENCES

Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect, 21*(1), 46-66.

Cushing, S. (1995a). *Fatal words: Communication clashes and aircraft crashes*. Chicago, IL: University of Chicago Press.

Cushing, S. (1995b). Pilot-Air Traffic Control communications: It's not (only) what you say, it's how you say it. *Flight Safety Digest, 14*(7), 47-53

Derwing, T. M. & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System, 30,* 155-166.

Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America, 93*, 1589-1608.

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. & Hillenbrand, J. (1984). Limits on pronunciation accuracy in adult foreign language speech production. *Journal of the Acoustical Society of America, 76*, 708-721.

Fraser, H. (1999). ESL pronunciation teaching: Could it be more effective? *Australian Language Matters, 7*(4), 7-8.

Fraser, H. (2000). Coordinating improvements in pronunciation teaching for adult learners of English as a second language. (Department of Education, Training and Youth Affairs, Australian National Training Authority ANTA Innovative Project) Canberra: Commonwealth of Australia.

Gass, S. M. & Varonis, E.M. (1985). Task variation and nonnative/nonnative negotiation of meaning. In S.M. Gass & C.G. Madden (Eds.), *Input in second language acquisition* (pp. 149-161). Rowley, MA: Newbury House Publishers.

Gilbert, J. B. (2005). *Clear speech: Pronunciation and listening comprehension in North American English. Teachers' resource book.* NY: Cambridge University Press.

Gilbert, J. B. (2010). Pronunciation as orphan: What can be done? *As We Speak: Newsletter of TESOL speaking, pronunciation, listening interest section, 7*(2).

Grant, L (2004). Demonstration presented at the 38th Annual TESOL Convention, Pre-Convention Institute: Essentials of Pronunciation, Long Beach, CA.

Grant, L. (2009). *Well said: Pronunciation for clear communication.* (3rd ed.). Boston, MA: Heinle & Heinle Publishers.

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

Hewings, M. (2004). *Pronunciation practice activities: A resource book for teaching pronunciatio*n. NY: Cambridge University Press.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*(1), 1-33.

Lardier, D. (1998). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research, 14*(4), 359-379.

Long, M. (2003). Stabilization and fossilization in interlanguage development. In C. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 4877-535). Malden, MA: Blackwell.

Long, M., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal, 82,* 357-71.

Major, R. (1995). Native and nonnative phonological representations. *IRAL, 33*(2), 109-127.

Major, R. (1998). Interlanguage phonetics and phonology: An introduction. *Studies in Second Language Acquisition*. *20*(2), 131-137.

Neufeld, G. (1977). Language learning ability in adults: A study on the acquisition of prosodic and articulatory features. *Working Papers in Bilingualism, 12,* 45-60.

Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition, 8,* 38-55.

Perkell, J. S., Matthies, M., Lane, H., Guenther, F., Wilhelms-Tricarico, R., Wozniak, J., & Guiod, P. (1997). Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models. *Speech Communication, 22*(2), 227-250.

Perkell, J. S., Guenther, F. H., Lane, H., Matthies, M. L., Perrier, P., Vick, J., Wilhelms-Tricarico, R. & Zandipour, M. (2000). A theory of speech motor control and supporting

data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics, 28*(3), 233-272.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 129-158.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed), *Speech Perception and Linguistic Experience: Issues in cross-language research* (pp. 3-45). Timonium, MD: York Press.

Vanderwood, M. L., & Nam, J. (2007). Response to intervention for English Language Learners: Current developments and future directions. In S. R. Jimmerson, M. K. Burns & A. M. VanDerHeyden (Eds.), *The Handbook of Response to Intervention: The Science and Practice of Assessment and Intervention* (pp. 408-417). NY: Springer.

Waldock, B. (1994, February 15). Major miscommunication mishaps. *Aviation Safety*, 3-5.

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.

PRONUNCIATION LEARNING STRATEGIES THAT IMPROVE ESL LEARNERS' LINKING

Veronica G. Sardegna, The University of Texas at Austin

This study explores the long-term effects of empowering 38 international graduate students with pronunciation learning strategies that they can use to improve their ability to link sounds within and across words. The effects of instruction are investigated through a mixed-method analysis that triangulates data from (a) a read-aloud test administered and recorded twice during the course, and again five months to two years after the course ended; (b) the same read-aloud test administered nine months after the testing in (a) to 23 of the 38 participants; (c) learners' responses to two questionnaires and a survey; and (d) the researcher's notes. The results reveal that students maintain a significant improvement over time regardless of their native language, gender, and length of stay in the US prior to instruction. There are, however, other learner characteristics and factors that seem to contribute to greater or lesser improvement over time. This paper identifies the pronunciation learning strategies taught for improving students' ability to link sounds, analyzes the factors contributing to lesser or greater improvement over time, and discusses the results and pedagogical implications of the study.

INTRODUCTION

This study explores empirical evidence in support of the Covert Rehearsal Model (CRM) (Dickerson, 1989, 1994, 2000; Hahn & Dickerson, 1999a, 1999b) for improving non-native students' (NNS) ability to link (or combine) sounds within words and at word boundaries without changing their phonetic qualities, as in (1), (2), (3), and (4), or by inserting a brief /j/ or /w/ sound between the sounds, as in (5):

(1) Type 1 – Consonant-to-Vowel Linking: *an error; is awesome; give in*
(2) Type 2 – Consonant-to-Same-Consonant Linking: *some music; Sue's snake;*
(3) Type 3 – Consonant-Stop-to-Other-Consonant-Stop Linking: *enthusiastic dad*; *adept*
(4) Type 4 – Consonant-to-Similar-Consonant[1] Linking: *come back; improve further*
(5) Type 5 – Vowel-to-Vowel Linking: *so exciting; diagonal; go in; play out*

CRM is a recursive process that comprises six key components (Dickerson, 2000):

1. Privacy
2. Oral practice
3. Speech monitoring

[1] Similar consonants are two adjacent consonants made in the same, or nearly the same, place in the mouth. When linked correctly, there is a change of articulation or voicing, but no break (as defined in Hahn & Dickerson, 1999a).

4. Comparing the performance with other models
5. Changing the performance to match the models
6. Practicing the changed performance aloud until fluent

CRM's goals are not only to improve students' ability to hear and articulate English sounds, rhythm, and melody, but also to improve their ability to predict the sounds, rhythm, and melody in words and phrases through the use of orthographically motivated rules and strategies (Hahn & Dickerson, 1999a, 1999b), such as those in (1)-(5) above for improving linking. CRM's instructional goals adhere to the goals of second language strategy research: Learning will be facilitated if students are more aware of and proficient in the use of a broad range of strategies that they can consciously select and use during language learning (Cohen, 1998; Weaver & Cohen, 1994).

Teachers implementing CRM, just as those pursuing strategy instruction (e.g., Chamot, Barnhardt, El-Dinary & Robbins, 1999; Oxford, 1996), explicitly teach students how, when, and why strategies can be used. During private practice, students learn to monitor and evaluate the relative effectiveness of their strategy use, and more fully develop their L2 comprehension, their problem solving skills, and their repertoire of useful strategies for pronunciation improvement. The kind of language produced in covert rehearsal is rehearsed speech focused on form. It is not spontaneous production. While engaged in CRM, students can practice aloud, use different pronunciation strategies, and monitor rule-use without distraction. In this way, CRM promotes students' autonomy and self direction, and empowers them to improve on their own after instruction ends.

RELEVANT LITERATURE

Despite arguments supporting instruction on both prosodic features and connected speech phenomena for improving NNS' fluency and intelligibility (Anderson-Hsieh, 1990; Brown & Kondo-Brown, 2006; Celce-Murcia, Brinton, & Goodwin, 2010; Dauer, 1992; Morley, 1994; Pennington & Richards, 1986), research evidence supporting such instruction has mostly focused on suprasegmentals like English stress and intonation (e.g., Benrabah, 1997; Derwing & Munro, 1997; Derwing, Munro & Wiebe, 1998; Field, 2005; Hahn, 2004; Kang, 2010; Pennington & Ellis, 2000; Sardegna, 2006, 2008, 2009). The few research studies that have investigated linking in the connected speech of NS and NNS suggest that linking is a marker of fluent speech (Alameen, 2007; Anderson-Hsieh, Riney, & Koehler, 1994; Hieke, 1984) and that native speakers' (NS) linking affects NNS' listening comprehension (Henrichsen, 1984). Unfortunately, very little is known about what affects the acquisition of linking and whether it can be learnable through explicit instruction.

Without studies investigating the effects of receiving instruction on linking, it is uncertain whether improved linking through instruction improves NNS' intelligibility or their listening comprehension of NS speech. Yet, despite lack of support from learnability studies, many English pronunciation textbooks teach linking through pronunciation rules and strategies (e.g., Celce-Murcia et al., 2010; Dauer, 1993; Gilbert, 1993; Grant, 2001; Hahn & Dickerson, 1999a; Hewings & Goldstein, 1998).

Studies that have looked at pronunciation strategies have mostly investigated learner's choices of strategies without prior instruction (Eckstein, 2007; Osburne, 2003; Peterson, 2000). In a different approach, Sardegna (2009) explored the long-term effectiveness of teaching students pronunciation strategies to improve English stress privately, and found strong evidence supporting instruction on CRM: over time participants improved significantly their ability to stress words, constructions, and phrases. The study also identified factors that may contribute to greater or lesser improvement with English stress: students' entering proficiency levels and degree of improvement with English stress during the course; their motivations and urgent need to improve their English stress; their prioritization to practice English stress in private; and the quality and, especially, the frequency with which they engaged in covert practice.

The present study answers calls for further research on the long-term effectiveness of pronunciation instruction (Derwing & Munro, 2005; Sardegna, 2009) and strategy instruction (Cohen & Macaro, 2007; Rees-Miller, 1993; Rubin, Chamot, Harris, & Anderson, 2007; Sardegna, 2009). It also answers calls for establishing methods to teach linking to non-native speakers (Alameen, 2007). More specifically, this study attempts to explore CRM's efficacy for teaching linking, and identify the factors that may facilitate or hinder its acquisition.

As this study examines students' improvement with linking while reading aloud, it makes no claims about their improvement in free speech. The analysis is therefore limited to three specific research questions:

1. What are the *short-term* effects of teaching ESL students how to improve linking during covert rehearsal as demonstrated in their reading of a passage at the beginning and end a four-month pronunciation course?
2. What are the *long-term* effects of teaching ESL students how to improve linking during covert rehearsal as demonstrated in their reading of a passage before instruction and from 5 to 38 months after instruction?
3. What factors contribute to greater or lesser improvement over time?

METHOD

Characteristics of the Pronunciation Course: The Teaching Intervention

The course used in this study was for international students at an American university. It taught students how to use pronunciation learning strategies to improve their oral English outside of class. It met for fifty minutes three times a week for four months. The materials, activities, and pronunciation rules given to students followed Dickerson's (1989; Hahn & Dickerson, 1999a) Covert Rehearsal Model (CRM). The basic premise of the model is that teachers should aim at equipping students with the predictive skills, pronunciation rules, and strategies that they need to work on the accuracy of their speech in private.

During the course, students learned a wide repertoire of pronunciation strategies. They also received written and oral feedback both during regular classroom activities and in five thirty-minute private meetings with the instructor. Both the training and feedback focused on improving students' most problematic pronunciation features as identified through a test

administered at the beginning of the course. The targeted features included segmental, suprasegmental, and connected speech features. Therefore, linking was just one of the instructional targets.

To improve their oral skills, students could use any combination of prediction, production, and perception strategies, either simultaneously or in sequence, while they worked under each of the six conditions for practice out of class delineated by CRM. For example, students could use the following combination of strategies for practicing linking:

1. *Privacy*: Find a place to work in private (no strategies involved).

2. *Oral Practice*: Retrieve stored information on how to divide long sentences into meaningful messages units, and on how to link the sounds within those units (remember not to link sounds across message units). Analyze the spelling and apply orthography-based prediction rules for linking (e.g., those listed on Hahn & Dickerson, 1999a, pp. 48-54) to figure out which sounds you should link and how you should pronounce them. Then read the predicted linked sounds aloud for smoothness of articulation.

3. *Speech Monitoring*: Evaluate the accuracy and fluency of your production as you read the phrases aloud many times, use backward build-ups, and apply visual guides (such as a line linking the sounds) to help you focus on the targeted sounds for linking.

4. *Comparing the Performance With Other Models*: Compare your production to a recording of the text, other sources that illustrate the same linking sound(s), or your own models from memory.

5. *Changing the Performance to Match the Models*: After consulting those sources/models, make all the necessary changes, and repair utterances or predictions to match the models.

6. *Practicing the Changed Performance Aloud until Fluent*: Practice the changes by reading aloud the phrases a number of times; and monitor your performance while you practice.

This list of strategies is by no means complete or exhaustive, but rather illustrative of the choices learners have. Students could use these cognitive and metacognitive strategies, simultaneously or in sequence, to predict, produce, and listen for one or more targets. In fact, the output of the process can, and often does, serve as the input to another round of the process.

Participants

Thirty-eight international graduate students (16 females and 22 males) from eight sections of this pronunciation course volunteered to participate in the study. All sections were taught by the same instructor and followed the same syllabus. The participants' ages ranged from 22 to 47 years old. Their native languages were: Chinese (19), Vietnamese (6), Korean (4), Thai (3), Turkish (3), French (1), Portuguese (1), and Spanish (1). The university had made the course a

graduation requirement for 22 of them due to their poor oral performance on the university's ESL Placement Test (EPT). The EPT consisted of an oral exam in which students talked about a topic and read some materials. Of the remaining 16 participants, half took the course because the EPT results recommended them to do so, while the other half took it for personal reasons. A total of 21 participants took the course on their first semester of graduate studies in the US, and eight on their second semester. The remaining nine had not been in the US for more than two years.

Experimental Design

The effects of instruction were investigated through a mixed method analysis that triangulated data from (a) pronunciation accuracy scores from a read-aloud test taken three times by all participants in a period from 18 to 29 months, (b) pronunciation accuracy scores of the same test in (a) taken nine months after the third test by only 23 of the 38 participants, (c) responses to two questionnaires and a survey, and (d) the researcher's notes on participants' comments about their practice in covert rehearsal.

Quantitative data were accuracy scores with respect to linking sounds within and across words from a pre-test (Test #1) and three post-tests (Test #2, Test #3, and Test #4). Using the same test four times allowed for an objective measure of students' progress over time. During the test, participants were told to monitor their production of *all* the target areas studied during the course, and not just focus on linking sounds correctly. Given the duration of the test (15-20 minutes), the number and variety of targets under assessment (segmental, suprasegmental, and connected speech features), and the time that passed between the tests (depending on the student and testing time, from 4 to 25 months), it seemed unlikely that the students would remember the words or sentences in the materials, and if they did, they had no way of knowing what targets were in focus for assessment and in what words or sentences. For example, linking targets were randomly identified for assessment forty times in a long passage, but the materials included not only the passage, but also a list of 22 words and 6 dialogs with other features targeted for assessment. Once a student finished reading the materials aloud, the percentage of incorrect verbal responses for each linking target marked on the rater's template was calculated by dividing the frequency of ticked responses by the total number of targeted instances. The result of this calculation constituted the participant's score for linking on that test.

All participants took Test #1 at the beginning of the course (Time 1/T1), Test #2 at the end of the course (Time 2/T2), and Test #3 from 5 months to 25 months after the course ended (Time 3/T3). The time between T2 and T3 was purposely varied to evaluate whether the variable "length of time after the course" significantly affected student progress. Participants were divided in three groups according to that variable: Group 1 = 5 months, Group 2 = 9 months, and Group 3 = 13-25 months. Only 23 of the participants took Test #4 nine months after Test #3 (Time 4/T4) because the others either graduated before it was administered, or withdrew from the study after T3. The time that passed between T3 and T4 was constant (9 months) because the purpose of Test #4 was to assess whether the trend observed from T2 to T3—forward or backward progress, or maintenance of performance—changed from T3 to T4. The percentage of participants that took Test #4 (60.5%) was considered large enough to generalize findings to all the participants in the study and to other studies of this kind. Table 1 shows the number of students and the time that passed between tests for each group.

Table 1. *Months Between Tests for Each Group*

Groups	N	Test #1 to Test #2	Test #2 to Test #3	Test #3 to Test #4	Total
1	9	4 months[a]	5 months	9 months	18 months
2	18	4 months	9 months	9 months	22 months
3	11	4 months	Over 13 months (13-25 months)	9 months	26-38 months

[a] Duration of the pronunciation course.

Qualitative data were collected through two questionnaires and a self-report survey that all participants completed at T3 and T4. The questionnaires elicited information about their practice in covert rehearsal, and personal information, such as their native language, age, gender, and length of stay in the US prior to instruction. The self-report survey gathered detailed information about students' use of pronunciation strategies in the intervening months. Students had to identify the strategies used and specify how much they had used them by choosing a number from 1 to 5 in a Likert scale. To minimize the limitations of self-reported data (as reported in White, Schramm, & Chamot, 2007), the information gathered through these methods was triangulated with participants' test scores, and their oral comments during testing times.

RESULTS

Improvement was calculated against a 100 percent base. For example, students' short-term improvement was calculated by subtracting their T1 mean percentage scores from their T2 mean percentage scores, and so on. A repeated measures ANOVA test was used to examine the overall time effect on students' performance from T1 to T3. This test measures participants' scores across each of the testing times and indicates whether their overall means differ significantly. The results revealed that students' performance differed significantly (*df-within* = 2; *df-error* = 74; $F = 111.866$; $p = 0.000$). Pairwise comparisons were computed to determine which mean differences were significant. Table 2 provides the mean percentage scores, standard deviations, and observed maximum and minimum scores for T1, T2, and T3 for the 38 participants. Table 3 shows the results of the pairwise comparisons. These comparisons revealed that participants' test scores increased significantly (30.99 percent, $p = 0.000$) during the course and, although their scores decreased significantly (-8.16 percent, $p = 0.000$) by T3, there is still a significant change (22.83 percent, $p = 0.000$) from T1 to T3 in students' performance (see student progress in Figure 1).

Table 2. *Descriptive Statistics for the Mean Percentage Scores for all Participants at T1, T2, and T3*

Test	N	M	SD	Minimum	Maximum
Test #1	38	52.171	13.964	12.5	75
Test #2	38	83.158	9.403	65	95
Test #3	38	75.000	12.053	42.5	95

Table 3. *Pairwise Comparisons for Linking for 38 Participants that Took Test #1, #2, and #3*

Pair	Difference	Pair 1	Pair 2	*p*-value[a]
Test #1-Test #2	-30.987*	52.171	83.158	0.000
Test #2-Test #3	8.158*	83.158	75.000	0.000
Test #1-Test #3	-22.829*	52.171	75.000	0.000

*The mean difference is significant at the 0.05 level. [a]Adjustment for multiple comparisons: Bonferroni

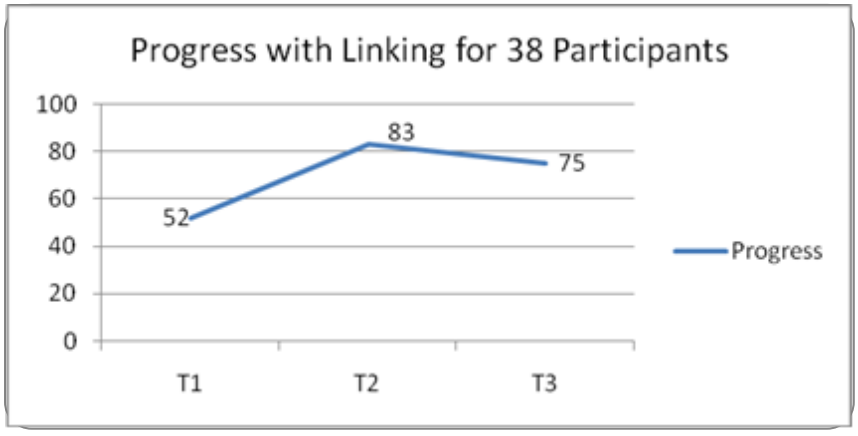


Figure 1. Mean Percentage Scores at T1, T2, and T3 for 38 participants.

Because change in accuracy levels from T1 to T2, and from T2 to T3 could result from a combination of factors unrelated to instruction and/or practice in covert rehearsal during and after the course, two Fixed Effects Models were computed to detect any possible variables affecting progress from T1 to T2 and from T2 to T3. Fixed Effects Models were preferred for these computations to ANOVA because these models do not assume balanced designs like

ANOVA. The factors possibly affecting T1-T2 progress were: native language, gender, length of stay in US prior to instruction, and accuracy level for linking at T1. The model revealed that the only significant factor was accuracy level for linking at T1 [$F(1,26) = 49.201$, $p = 0.000$, Table 4]. This factor had a strong negative correlation with degree of progress between T1 and T2 ($r = -0.770$ at $p = 0.000$), indicating that the higher the level of accuracy at T1, the smaller the improvement achieved from T1 to T2.

Table 4. *Type III Tests of Fixed Effects for Improvement in Linking From T1 to T2*

Source	Numerator *df*	Denominator *df*	*F*	*p*-value
Intercept	1	26	124.669	0.000
Language	7	26	1.676	0.159
Gender	1	26	0.091	0.766
Length of Stay in US Prior to Instruction	2	26	0.606	0.553
Accuracy Level for linking at T1	1	26	49.201	0.000*

Dependent Variable: Progress from T1 to T2. *Significant at the .05 level.

The Fixed Effects Model to detect variables affecting progress from T2 to T3 included three factors–group (i.e., length of time between T2 and T3), language, and gender,–and two covariates–progress from T1 to T2, and practice with linking (i.e., amount of practice in covert rehearsal[2]). Despite the differences in testing time at T3 for the three groups, the test detected no significant difference among the groups [$F(2,25) = 1.383$, $p = 0.269$]. We therefore conclude that length of time between T2 and T3 did not significantly affect students' scores at T3 (see Table 5).

Table 5. *Type III Tests of Fixed Effects for Improvement in Linking From T2 to T3*

Source	Numerator *df*	Denominator *df*	*F*	*p*-value
Intercept	1	25	1.990	0.171
Group	2	25	1.383	0.269
Language	7	25	0.954	0.485
Gender	1	25	0.050	0.826
Progress During the Course	1	25	5.940	0.022*
Practice with Linking	1	25	6.817	0.015*

Dependent Variable: Progress from T2 to T3. *Significant at the 0.05 level.

[2] This amount corresponds to the sum of activities students reported doing to improve linking from T2 to T3 in their self-reported survey at T3.

In fact, the model detected significant differences with two other variables: (a) progress during the course (that is, percentage of improvement with linking from T1 to T2) [$F(1,25) = 5.94$, $p = 0.022$], and (b) students' amount of practice with linking from T2 to T3 [$F(1,25) = 6.817$, $p = 0.015$]. The estimate of these covariance parameters was significant ($p = 0.000$) and their association with T2-T3 progress was large enough for two separate but related variables: $r = -0.427$ at $p = 0.004$ for T1-T2 progress, and $r = 0.503$ at $p = 0.001$ for practice with linking. T1-T2 progress negatively correlated with T2-T3 progress, indicating that the higher the percentage of increase in accuracy from T1 to T2, the higher the percentage of decrease in accuracy from T2 to T3. Practice with linking positively correlated with T2-T3 progress, indicating that the more students practiced, the greater their improvement.

With Group having no significant effect on student performance at T3, it was possible to compare means across testing times for the group as a whole. This comparison revealed a clear group trend (Figure 1): significant improvement during and after the course both regardless of the time that passed after T2 and despite the significant decrease in accuracy from T2 to T3. This trend was evident in 30 of the 38 participants. Only eight participants increased their accuracy level from T2 to T3: S#5 by 2.5 percent; S#1, S#2, S#18 by 5 percent; S#22 and S#36 by 7.5 percent; S#24 by 10 percent; and S#25 by 15 percent. The others, with rare exceptions, all decreased generally but by not more than 10 percent. Besides amount of practice, students with higher improvement or maintenance of improvement at T3 reported frequency of practice, high internal/external motivations to improve, such as aspiring to get financial aid with better pronunciation, and prioritization of practice with linking over other targets. For example, S#24, who improved 10 percent from T2 to T3, reported "*I practice once a week at home. I read the newspaper aloud."* The researcher also observed that some students who did not improve as much despite having reported large amounts of practice were in fact not practicing in the manner instructed during the course. The quality of their practice mattered, too.

To test maintenance of improvement, 23 of the 38 participants (60.5 percent) were tested a fourth time nine months after T3. A repeated measures ANOVA test was used to examine the overall time effect on students' performance from T1 to T4 for the students who took all four tests. The results revealed that students' performance differed significantly over time (*df-within* = 3; *df-error* = 66; $F = 56.887$; $p = 0.000$). Pairwise comparisons were also computed to determine which mean differences were significant (Table 6).

Table 6. *Pairwise Comparisons for Linking for the 23 Students that Took Tests #1, #2, #3, and #4*

Pair	Difference	Pair 1	Pair 2	[a]*p*-value
Test#1-Test #2	-32.826*	49.783	82.609	0.000
Test #2-Test #3	9.783*	82.609	72.827	0.001
Test #3-Test #4	-0.3261	72.827	73.152	0.890
Test #1-Test #3	-23.043*	49.783	72.827	0.000
Test #1-Test #4	-23.369*	49.783	73.152	0.000

*The mean difference is significant at the 0.05 level. [a]Adjustment for multiple comparisons: Bonferroni.

These results indicate a clear trend in students' performance over time for improving linking: a large and significant improvement from T1 to T2 (around 33 percent, $p = 0.000$), a significant decrease after intensive instruction from T2 to T3 (around 10 percent, $p = 0$ .001), and a plateau from T3 to T4 ($p = 0.890$). Despite students' significant decrease in accuracy from T2 to T3, they maintained a significant change from T1 to T3 (23 percent, $p = 0.000$) and from T1 to T4 (23 percent, $p = 0.000$). In fact, there was almost no difference in progress between T1 and T3, and T1 and T4. Figure 2 shows students' overall performance from T1 to T4.

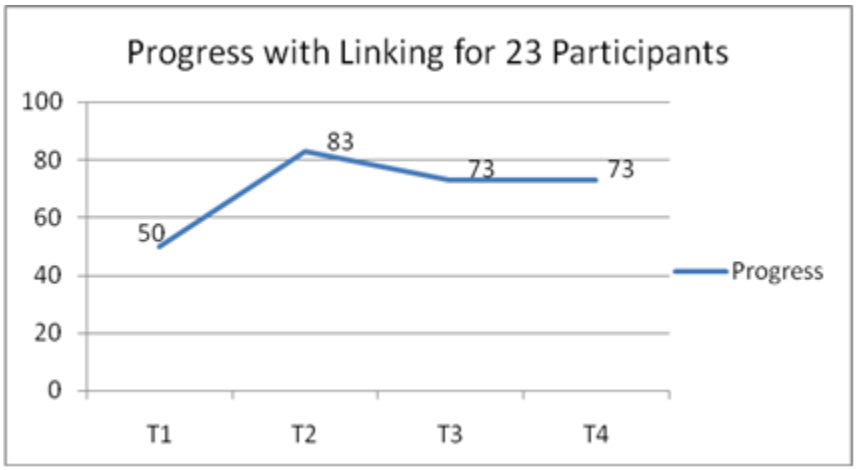


Figure 2. Mean Percentage Scores at T1, T2, T3, and T4 for 23 participants.

DISCUSSION

The evidence from the scores at T1 and T2 as a group (31 percent increase for 38 participants; 33 percent increase for 23 participants) is strong enough to answer research question #1 in the affirmative. Participants clearly showed short-term improvement with linking while reading aloud (Tables 3-6). After four months of instruction on and practice with pronunciation learning strategies for improving linking, students' scores increased significantly. The evidence from the participants' scores from T1 to T3/T4 as a group is also strong enough to answer research question #2 in the affirmative: participants clearly showed long-term improvement with linking while reading aloud. Their scores increased significantly from T1 to T3/T4 (9-34 months after instruction) despite their decrease in accuracy from T2 to T3 (see Tables 3-6) and regardless of the time that passed between T2 and T3 (see Table 5).

In fact, a close look at students' individual scores from T1 to T2, and from T1 to T3 revealed that all participants improved short-term and long-term regardless of individual learner differences. The results indicate that the methodology used during the course may have facilitated such improvement. Like Derwing et al.'s (1998) study, these findings provide evidence in support of the effectiveness of pronunciation instruction. The results also support Dickerson's (1994) claim in favor of empowering students with explicit pronunciation rules that they can use to self-correct and self-monitor their speech production. In addition, they extend the results reported in Sardegna (2009) with respect to English stress to linking. More research is needed to extend this result to other suprasegmentals and to segmentals. Also, to make a stronger claim on the efficacy of the methodology used, the results of instruction should be compared to other studies of the same kind utilizing other methodologies or with groups of students receiving no instruction.

This study's experimental design allowed testing for variables other than "receiving no instruction" which shed some light into the factors contributing to greater or lesser individual short and long-term progress with linking. Differences in improvement during instruction could not be predicted on the basis of students' gender, language background, or length of stay in the target language community before instruction, but it could be predicted on the basis of students' entering proficiencies with linking. Students with a higher level at the beginning of the course improved less than those with a lower level, which corroborates findings from research on other pronunciation features targeted during instruction (Dickerson, 2002; Sardegna, 2006, 2009). Hence, entering proficiencies seem to be strong indicators of students' ultimate accuracy levels at the end of the course.

Maintenance of learning after a course provides evidence that the short-term improvement achieved during intensive instruction is lasting and, therefore, effective (Beebe, 1988; Sardegna, 2009). The study corroborates claims (Beebe, 1988; Sardegna, 2008, 2009) that, after receiving pronunciation instruction, and improving as a result of that instruction, students generally decrease in accuracy, or as Beebe puts it, suffer phonological "backsliding." This trend, therefore, seems to be indicating how pronunciation development takes place: students tend to decrease their accuracy somewhat after increasing it during intensive instruction.

In addition, the study revealed two factors that may have contributed to differences in individual improvement after the course: practice in covert rehearsal and T1-T2 improvement. The more students practiced after the course, the more they improved or maintained the level of

improvement achieved during the course; the higher their improvement during the course, the more they decreased in accuracy at T3. It appears that before students reached a plateau (observed from T3 to T4), they decreased in accuracy, and that decrease tended to be larger if there had been a high increase in accuracy during the course, and especially for those students who reported no or little practice after the course ended. Students' comments revealed three other factors that may have contributed to student differences from T2 to T3: students' frequency and quality of practice, their motivations to improve, and their choice of linking for focused practice.

With this information, we can now answer research question #3: What factors contribute to greater or lesser improvement over time? The factors that may contribute to greater or lesser improvement with linking are the following:

1. Entering proficiency level with linking.
2. Degree of improvement with linking during the course.
3. Quantity, quality, and especially frequency of practice with linking in covert rehearsal.
4. Strong intrinsic and extrinsic motivations to improve.
5. Prioritization of linking over other targets for focused practice.

These factors closely match those reported in Sardegna (2009) for long-term improvement with English stress. It would be interesting to investigate whether these factors also affect long-term improvement with segmentals as well as with suprasegmentals other than English stress.

Also, just like with English stress, students' production stabilized a few months after the course. Students reached a new plateau (see T3 and T4 mean scores in Figure 2). This observation is another indicator that it was the course that was effective in producing a long-lasting effect on students' linking, and not just the students' willingness to improve. In as little as five months after instruction, students' focused practice or motivations did not seem to make much difference, which would explain why the different testing times between T2 and T3 did not affect much the overall group trend, or students' individual scores. As this study shows, to make long-lasting improvements, students first need to improve significantly (which they can only do through intensive instruction). Decrease in accuracy after the course, followed by a plateau, is to be expected. Yet, what students do immediately after the course seems to influence the expected initial decrease.

PEDAGOGICAL IMPLICATIONS

Pronunciation progress is gradual, and does not happen overnight (Dickerson, 1989). It depends on students' considerable commitment of time and effort to improve (Acton, 1984; Murphy, 1991). This study shows that teaching for empowerment provides students the resources they need to improve linking on their own during and after the course. Improvement using this methodology is lasting and effective. It puts the burden on the students, and makes them responsible for their own progress.

However, teachers should be realistic about what to expect from a pronunciation course. Students do not generally progress much on their own after the course ends. In fact, they generally decrease in accuracy by about 10 percent, or more if there was a high increase (e.g., more than 30 percent increase) from T1 to T2. However, if their increase in accuracy during the course is high, they have a high chance of maintaining that accuracy (and not decreasing much)

if they continue their practice in covert rehearsal in the few months following the course. Frequent practice in covert rehearsal makes students focus on forms, systematically, and with accurate rules and models, thereby consolidating the strategies, sound articulations, and melody of English. This practice seems to help students slowly internalize the rules, thereby requiring less processing and practice time. If that is the case, then accuracy changes evidenced in students' ability to stress and link sounds should also eventually become evident in extemporaneous speech. A fruitful avenue for future research would be to test this hypothesis.

It is also hypothesized that, if teachers control somewhat the factors contributing to greater improvement, students may have higher chances of success.

1. Through constant feedback, a teacher can increase students' motivations to work in covert rehearsal. Recognition of students' achievements may increase their sense of self-efficacy (Zimmerman, 2000) and their willingness to continue working on their own after the course ends.

2. A teacher can gauge students' prioritizing and focused practice by providing them with frequent assessments. Students that finish the course knowing what areas they still need to improve are more likely to prioritize practice in those areas after the course ends.

3. A teacher can develop materials that increase quality, quantity, and frequency of practice, and post them online. Students are more likely to continue using the same materials after the course ends if these are readily available for their practice in private.

CONCLUSION

This study provides evidence in favor of CRM for improving students' ability to link sounds within and across words while they read aloud. Students significantly improved over time despite an initial decrease in accuracy after the course ended. Five factors were found to contribute to greater or lesser improvement over time:

- Entering proficiency level with linking
- Degree of improvement during the course
- Quantity, quality, and especially frequency of practice in covert rehearsal after instruction
- Strong intrinsic and extrinsic motivations to improve
- Prioritization of one target over another for focused practice

Teachers can increase students' chances of long-term pronunciation improvement by providing the resources they need to work on their own, and helping them develop a sense of self-efficacy. Students who believe that they can positively change their pronunciation on their own are more likely to engage in covert practice and improve as a result.

ABOUT THE AUTHOR

Veronica G. Sardegna is an Assistant Professor at The University of Texas at Austin where she teaches courses on methodology, pronunciation, and second language acquisition for both the MA and doctoral programs in Foreign Language Education. Her current research work focuses on the effectiveness of using language learning strategies and instructional technology for improving ESL/EFL students' pronunciation and business writing skills. Her most recent publication is *The Practiced Business Writer: An ESL/EFL Handbook* (2009), published by Briefings Media Group. Contact her at sardegna@mail.utexas.edu.

REFERENCES

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly, 18,* 71-85.

Alameen, G. (2007). *The use of linking by native and non-native speakers of American English.* Unpublished MA thesis, Iowa State University.

Anderson-Hsieh, J. (1990). Teaching suprasegmentals to international teaching assistants using field-specific materials. *English for Specific Purposes, 9,* 195-214.

Anderson-Hsieh, J., Riney, T., & Koehler, K. (1994). Connected speech modifications in the English of Japanese ESL learners. *Ideal, 7,* 31-52.

Beebe, L. M. (1988). The sociolinguistic perspective. In L. M. Beebe (Ed.), *Issues in second language acquisition: Multiple perspectives.* New York: Newbury House.

Benrabah, M. (1997). Word-stress: A source of unintelligibility in English. *IRAL, 35,* 157-165.

Brown, J. D., & Kondo-Brown, K. (eds.). (2006). *Perspectives on teaching connected speech to second language learners.* Honolulu, HI: University of Hawai'I at Manoa, National Foreign Language Resource Center.

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., Barnhardt, S., El-Dinary, P. B., & Robbins, J. (1999). *The learning strategies handbook.* White Plains, NY: Addison-Wesley Longman.

Cohen, A. D. (1998). *Strategies in learning and using a second language*. Harlow, England: Longman.

Cohen, A. D. & Macaro, E. (2007). *Language learner strategies.* Oxford, UK: Oxford University Press.

Dauer, R. M. (1992 March). *Teaching the pronunciation of connected speech*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, Vancouver, British Columbia, Canada. Retrieved from ERIC database (ED 354 777)

Dauer, R. M. (1993). *Accurate English: A complete course in pronunciation.* Prentice Hall Regents.

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

Derwing, T. M., Munro, M. J, & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48,* 393-410.

Dickerson, W. B. (1989). *Stress in the speech stream: The rhythm of spoken English*. [Student and Teacher's Text]. Urbana: University of Illinois Press.

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.

Dickerson, W. B. (2000 March). *Covert rehearsal as a bridge to accurate fluency*. Paper presented at International TESOL, Vancouver, BC, Canada.

Dickerson, W. B. (2002). Grading pronunciation progress: The roles of absolute and relative scales. In W. Davis, J. Smith, & R. Smith (Eds.), *Ready to Teach* (pp. 59-166). New Forums Press, Inc..

Eckstein, G. T. (2007). *A correlation of pronunciation learning strategies with spontaneous English pronunciation of adult ESL learners.* Unpublished MA Thesis, Brigham Young University.

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

Gilbert, J. B. (1993). *Clear speech: Pronunciation and listening comprehension in North American English.* Student's book (2nd ed.). Cambridge: Cambridge University Press.

Grant, L. (2001). *Well Said: Pronunciation for clear communication* (2nd ed.). Boston: Heinle & Heinle.

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

Hahn, L., & Dickerson, W. (1999a). *Speechcraft: Discourse pronunciation for advanced learners.* Ann Arbor: University of Michigan Press.

Hahn, L., & Dickerson, W. (1999b). *Speechcraft: Workbook for academic discourse.* Ann Arbor: University of Michigan Press.

Henrichsen, L. E. (1984). Sandhi-variation: A filter of input for learners of ESL. *Language Learning, 34,* 103-126.

Hewings, M., & Goldstein, S. (1998). *Pronunciation plus*. Cambridge: Cambridge University Press.

Hieke, A. E. (1984). Linking as a marker of fluent speech. *Language and Speech, 27,* 243-354.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System, 38,* 301-315.

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.

Murphy, J. (1991). Oral communication in TESOL: Integrating speaking, listening and pronunciation. *TESOL Quarterly, 1,* 51-75.

Osburne, A. G. (2003). Pronunciation strategies of advanced ESOL learners. *IRAL, 41,* 131-143.

Oxford, R. L. (1996). *Language learning strategies around the world: Cross-cultural perspectives* (Tech. Rep. No. 13). Honolulu, HI: Second Language Teaching and Curriculum Center, University of Hawaii Press.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal, 84,* 372-389.

Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly, 20,* 207-225.

Peterson, S. S. (2000). *Pronunciation learning strategies: A first look.* Retrieved from ERIC Database (ED 450 599).

Rees-Miller, J. (1993). A critical appraisal of learner training: Theoretical bases and teaching implications. *TESOL Quarterly, 27,* 679-689.

Rubin, J., Chamot, A. U., Harris, V., and Anderson, N. J. (2007). Intervening in the use of strategies. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. 141-164). Oxford, UK: Oxford University Press.

Sardegna, V. G. (2006 March). *Achieving pronunciation accuracy through covert rehearsal.* Paper presented at the 40th Annual TESOL Convention and Exhibit, Tampa, FL.

Sardegna, V. G. (2008 April). *Pronunciation learning strategies that work.* Paper presented at the 42nd Annual TESOL Convention and Exhibit, New York, NY.

Sardegna, V. G. (2009). *Improving English stress through pronunciation learning strategies.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign (UMI No. 3363085).

White, C., Schramm, K., & Chamot, A. U. (2007). Research methods in strategy research: Re-examining the toolbox. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. 93-116). Oxford, UK: Oxford University Press.

Weaver, S. J., & Cohen, A. D. (1994). Making learning strategy instruction a reality in the foreign language curriculum. In C. Klee (Ed.), *Faces in a crowd: The individual learner in multisection courses* (pp. 285-323). Boston: Heinle & Heinle.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25,* 82-91.

Huffman, S. (2011). Using mobile technologies for synchronous CMC to develop L2 oral proficiency. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 122-129), Ames, IA: Iowa State University.

USING MOBILE TECHNOLOGIES FOR SYNCHRONOUS CMC TO DEVELOP L2 ORAL PROFICIENCY

Sarah Huffman*, *Iowa State University

As the use of the mobile technology grows increasingly ubiquitous among populations worldwide, it is essential that second language instructors, researchers and curriculum designers understand their full functional capacity before implementing mobile technologies into activities and assignments in second language classroom. This paper explores the range of potential for multimedia mobile technologies in developing L2 oral proficiency and focuses on synchronous communication in particular. The paper also addresses points to consider before implementing mobile technologies into L2 oral language instruction and makes recommendations for future research on the use of these technologies in developing oral, and other, language skills. Recognition of effective and appropriate uses for mobile technologies in synchronous CMC environments will help facilitate meaningful and effectual language learning and enable learning environments to transcend the physical boundaries of the classroom.

INTRODUCTION

As the world and the classroom undergo a shift "from print to post-print text cultures" it becomes increasingly challenging for language learning professionals to delineate concepts of literacy, proficiency and competency across linguistic skill sets (Lankshear, Gee, Knobel, & Searle, 1997, p. 3). Improving second language learners' oral skills in their L2 is beginning to require an increasingly multidimensional approach with persistent consideration of learners' capacities and contexts of learning.

Recently, a growing number of second language teachers and researchers have begun implementing synchronous CMC (Computer Mediated Communication) tools in the CALL (Computer Assisted Language Learning) classrooms devoted to developing L2 oral proficiency. Synchronous CMC tools allow for real-time communication across the skill sets of reading, writing, speaking and listening and it is expected that investigations will swell as new technologies allowing for synchronous CMC are introduced (Thorne, 2008). One of the most popular synchronous CMC technology, one that has transformed how teachers and students view the learning process, is mobile technology. These "mobile, wearable and pervasive technologies" allow language learning to become integrated into the current of students' everyday lives as learners' real and virtual environments merge (Hampel & Hauck, 2006, p. 16).

Mobile technologies are gaining popularity in developing L2 oral proficiency for a number of reasons. First, language learning in a second language context is not always a viable or realistic option for most language learners. The fact that many want or need to learn a second language, but not all learners are able to engage in face-to-face tutorials in a second language context has given way to commercial and academic institutions offering distance education courses for second language learning. Because distance education depends so heavily on technology, "appropriate technological media are a precondition of improved language learning in distance mode" (Wang, 2004, p. 374).

SYNCHRONOUS MOBILE TECHNOLOGIES USED IN DEVELOPING L2 ORAL PROFICIENCY

Products and Tools

When many think of mobile technology, cell phones typically come to mind. Multimedia cell phones, mobile phones equipped with audio and sometimes video recording devices, enable recording of students' oral speech. Such devices are of particular interest to second language instructors who wish to have students practice their oral language skills in in-class or at-home activities. Also, because multimedia phones are Internet accessible, voice emailing, Internet relay chat, instant messaging, audio conferencing and real-time voice chat are just a few of the means by which learners are allowed to interact in synchronous communication with other NNSs or NSs to practice their L2 oral language.

Another mobile technology that allows for learners to practice their L2 oral skills is the portable mp3 player. Mp3 players , like the iPod, one of the most popular mp3 players, allow for recording of digital audio files, accessing and manipulation of audio and video content and provide high-quality resolution in a compressed format that makes files easy to share. In asynchronous CMC, teachers may use portable mp3 players to develop students' oral language skills by designing projects or assignments that require students to create an audio (or sometimes audio and video) file called a podcast.

PDAs (Personal Digital Assistants) with multimedia capabilities also hold the potential for use in developing L2 oral proficiency. In addition to having Internet access, most PDAs also function as cell phones. Smartphones, like the BlackBerry, and Android phones, like the Motorola droid, are well-known examples. Though PDAs are not as proliferate as multimedia cell phones and iPods, they still hold possibilities for building students' oral proficiency, as they have Internet access, available USB cables or Bluetooth to connect to the Internet when wireless is unavailable, and audio/visual recording capabilities.

Programs and Applications

There are numerous free and commercial Internet-based programs and applications that allow for students to use their mobile technologies to engage in synchronous communication online. Many of these have been developed by curriculum designers or educational institutions to fit the needs of facilitating NS-NNS communication in distance learning environments or in foreign language learning settings.

One of the most popular synchronous voice chat programs is Skype. Not only does Skype provide real-time oral communication between any speaker who has downloaded the free program, but also it allows for video streaming and written text chat in addition to the audio. This synchronous communication replaces the delayed email exchanges with real-time chat exchanges and allows second language learners to practice oral production skills via mobile devices with native speakers from across the globe (Volle, 2005). The video accessibility enhances the real-time audio feature by allowing additional nonverbal language support (facial expressions, gestures, body language), a feature language learners often remark as extremely useful in developing their L2 oral proficiency (Robin, 2007).

A software program called MobiLearn permits students with PDAs to practice by repeating provided audios of NS pronunciation of words through a multimedia translator function. These "talking phrasebooks" work as voice recognition software, in a sense, as they offer audio clips of NSs pronouncing a word, then give learners a chance to practice saying the word into the device (Chinnery, 2006, p. 11). Because learners are given instant feedback, they can notice gaps in their own speech and that of the NS and make necessary adjustments in their pronunciation.

Commercial videoconferencing software, such as NetMeeting (a program that can be downloaded for free for Windows), is an Internet-based videoconferencing application that was designed specifically to enhance interactive language learning in distance education courses (Wang, 2004). NetMeeting and other programs like VideoVoxPhone and CUseeMe enable synchronous audio, video, simultaneous written text chat, document sharing, file transfers, and some whiteboard resources to users, and are used chiefly by instructors and students in distance language learning situations. This commercial software, along with programs like *Lyceum*, a similar Internet-based audio-graphics conferencing program developed by researchers at Open University in the UK, let students, tutors and instructors hear and speak to one another while attending real-time virtual classrooms (Hampel & Hauck, 2006).

AFFORDANCES OF SYNCHRONOUS CMC-CAPABLE MOBILE TECHNOLOGIES IN PROMOTING ORAL PROFICIENCY

Because mobile technologies are "no longer part of the specialized landscape of the L2 learner," but rather, "make up the everyday L1 machine-mediated world" of communication, the tools hold a wide range of potential for use in the second language learning classroom (Robin, 2007, p. 109). Use of mobile technologies for language learning would simply require tapping a pre-existing resource in the lives of many students. The International Telecommunication Union (ITU) expects cell phone subscriptions to surpass 5 billion in 2010 (Whitney, 2010). To claim this number is significant in consideration of the overall world population (6.8 billion) would be an understatement. It also seems the explosion in subscriptions is occurring in both developed and developing countries, and among younger and younger populations (eligible subscribers comprise those 13 years and older). The ITU also expects the demand for mobile access to the Internet to increase beyond 1 billion mobile broadband subscriptions within the year (Whitney, 2010). Reassigning the usage of the technologies for purposes that would help develop L2 oral proficiency would merely be taking full advantage an already widespread language learning resource.

User-friendliness

Because the mobile technologies are often an integral part of a student's daily life and students are often quite familiar with manipulating their devices' tools, capabilities and settings, incorporating the technologies into the learning process makes the means of learning user-friendly (Wang, 2004). Students' familiarity with the interface, especially when they are using owned personal devices, facilitates greater ease in the learning process through increased comfort with the environment. The learning process is more personalized through students' closeness with the learning medium, a factor that develops greater self-confidence in a language learner (Norbrook & Scott, 2003); as a result of increased confidence and feelings of familiarity, students' may be more willing to take risks in their L2.

Convenience

The portability of mobile technologies is another factor which motivates both language learners and language teachers to use them (Norbrook & Scott, 2003). Mobile media are more handy (more lightweight and transportable) than desktop or laptop computers and, therefore, more conveniently accessed. Because many learners carry mobile devices with them every day, students have the freedom to study when and where they want. Likewise, teachers using mobile technologies are able to update assignments, access student work and interact with students at their convenience.

Accessibility

Mobile technologies not only provide convenience for the teacher and learner, but the anytime anywhere dimension of synchronous CMC has been shown to release pressure from students who many feel overwhelmed by the aural and oral demands of face-to-face communication, making communication

more accessible. Unlike face-to-face communication, real-time computer-mediated conversations allow for a delay in response; this delay, permitting an increase in learners' reaction time, puts less pressure on students to speak quickly (Payne & Whitney, 2002). Also, research has shown that synchronous CMC environments encourage quieter students to interact more (Warschauer, 1996; Kern, 1995). Hesitant or shy students are more willing to participate in synchronous CMC conversations, sometimes even more than students who tend to dominate discussion in the physical environment of the classroom, perhaps because they feel less pressure from their interlocutors in an online environment (Warschauer, 1996).

Accessibility is also extended by the technologies' capacity to broaden students' access to native speakers of the L2, samples of the L2 and activities in the L2. In this way, mobile devices are particularly useful in foreign language settings or distance learning settings where students cannot practice their L2 in a shared physical setting with NSs. Synchronous CMC allows for students to participate in spontaneous production of their L2 in interactions with NSs, contributing to greater L2 oral fluency (Wang & Sun, 2001; Wang, 2004). Students may hone their pragmatic and discourse competencies with native speakers as they repair, clarify or confirm messages (Bachman & Palmer, 1996). In his study of NNS-NS face-to-face interaction, Linnell (1995) found that clarification requests on the part of the NNSs sparked not only greater syntax modification by the NNSs, but also NNSs noticing the gap between their speech and the speech of the NSs. This immediate feedback, whether from a NS or an instructor, concerning learners' oral production is invaluable to the learner (Thorton & Houser, 2003).

Learner autonomy

In many ways, the language learner is empowered through the use of mobile technologies. Murphy cites that autonomy, independence and responsibility are among the greatest assets of using mobile technologies in language learning (2008). The pacing of activity completion or material access is dependent on the students; as Payne and Whitney (2002) note, "the notion that learners can practice speaking in an environment where affect and rate of speech are minimized is very appealing" (p.25). In terms of developing oral proficiency, the student participating in a mobile language learning environment is empowered by experiencing a mixture of self-directed or instructor-directed classes. In voice chat rooms, through Net-conferencing or through live editing programs, the student participates in classes that are largely student-centered and student-dependent, as the learner's output is mandatory for task completion.

Cost-effectiveness

Institutional costs for learning materials may also be curbed if instructors are able to have students use their owned mobile phones, PDAs or iPods. Utilizing the multimedia tools students already own is a cost-effective means of enhancing learners' overall language learning experience (Wang, 2004). Additionally, because some synchronous communication-based programs and applications can be downloaded for free from the Internet, real-time CMC (whether mobile or static) remains a feasible and functional fixture in the promotion of language proficiency.

LIMITATIONS OF MOBILE TECHNOLOGIES IN DEVELOPING ORAL PROFICIENCY IN SYNCHRONOUS CMC

Incorporation of synchronous CMC activities that require oral production from speakers of lower language proficiency should be carefully considered before being implemented in the classroom. As noted by several researchers (Heins, Duensin, Stickler & Batstone, 2007; Kiernan & Aizawa, 2004; Wang, 2004), beginning language learners are more reticent and tend not to initiate exchanges in real-time conversations, especially when the interactions include NSs of the L2. Beginners' limited linguistic abilities require increased reliance on structured L2 input and, in general, beginners should not be

expected to provide large amounts of L2 oral output independently or dependently of prompts (Want, 2004). Also, because "negotiation of meaning can take a longer time or can be difficult...when the learner's proficiency is low", instructors should be selective about how and how often they include synchronous CMC activities in their beginner-level classes (p. 381).

Another limitation is that oral synchronous CMC carries a tendency to have learners concentrate more on the rate of production rather than on the accuracy of the language. Wang (2004) remarks that oral interaction in real-time contexts "offers more spontaneity and fluency than written interaction, but accuracy may be at risk because students do not have time to prepare what they wish to say in a real-time situation" (p. 381). This lack of attention to accuracy, on the students' and teachers' parts, may be detrimental to language learners, especially in preliminary stages of their L2 oral language development. Instructors must pay equal attention to precision of the language as they do to the fluency of the production (Harley, 1993).

Mobile technologies also present a diminished capacity for displaying quality audiovisual materials that may supplement instruction of oral practice than laptop or desktop computers. The reduced bandwidth of mobile technologies, as compared to PCs, complicates learners' accessing, dissemination and production of quality video and audio, causing side effects like distortion and screen freezing which interrupt the flow of synchronous speech. These obstacles to real-time interaction are shown to frustrate learners in addition to causing miscomprehension or unintelligibility (Wang, 2004). Acknowledgement of these limitations when designing a synchronous CMC activity that uses mobile technologies is crucial so complications that may discourage the learner from producing her L2 may be avoided.

Another technological difficulty accompanying mobile technology deals with the reliability and availability of Internet access. The synchronous aspect of real-time interaction in the L2 requires mobile technologies to have high-speed Internet access through a wireless router or cell phone provider. Easy internet access and strong connections or signals are not always available to distance learners studying in their home countries. Likewise there is an unequal distribution of Internet access and mobile technological advancements across socioeconomic classes and geopolitical spheres (Van Dijk, 2005). Instructors' knowledge of their learners' Internet access and technological capabilities is fundamental to choosing appropriate tasks and media that will promote L2 oral proficiency by realistic means.

To avoid letting the technological aspect of mobile learning environments become the primary focus of class activities, researchers (Copaert, 2004; Salaberry, 2001) suggest that in mobile language learning environments, as with computer-mediated learning environments, it is essential to first develop the language learning environment before determining the role of mobile technologies in the classroom. Mobile technologies "are not in and of themselves instructors; rather, they are instructional tools" which require thoughtful application (Chinnery, 2006, p. 9). As the focus on technology intensifies in a language learning course, it remains integral to "focus on the learner ahead of the technology" (Chinnery, 2006, p. 9). These recommendations are reasonable considering the temporal and monetary investments involved in the use of mobile technologies in the classroom and in light of the lack of evidence proving them to be more effective than traditional second language learning means (Beatty, 2003).

CONCLUSION AND RECOMMENDATIONS FOR FUTURE RESEARCH

As mobile technologies continue to become embedded in the personal lives of language learners, how the devices can be used to cultivate rich social interactions should be a principle concern for second language learning professionals. Interaction is "indispensable to language learning for the simple reason that language itself is a means of communication and interaction" (Wang, 2004, p. 374); thus, promoting learners' interaction in the L2 through readily available and multimedia means holds enormous potential

for developing not only L2 oral proficiency, but also receptive and productive language skills on the whole.

Revising how instructors and researchers see learning environments is a necessary first step in incorporating mobile technologies, or other emerging technologies, into the second language learning classroom. Hampel and Hauck (2006) argue that it "is not sufficient to see the new learning spaces as replicates of conventional face-to-face settings," but rather as blossoming, equally genuine arenas for authentic communication (p. 3). Revising our framework of what constitutes *environment* may require a reconceptualization of virtual realities as authentic communicative settings.

Limitations of both mobile technologies and synchronous CMC environments must also be investigated before claiming their effectiveness in building L2 oral language skills. Conducting further research on how students of varying proficiency levels can handle real-time communicative tasks is necessary to ensure speakers of lower proficiency levels do not fall behind or become less interactive in mobile-learning CMC environments. Adaptations of activities suitable for advanced or intermediate level learners may be necessary before overwhelming beginning learners with overly complicated oral production tasks in real-time situations.

The advancement of current and development of new, more sophisticated multimedia technologies is also needed. Multimedia-capable mobile technologies afford "anyplace-anytime access to and production of Internet-distributed text, video and audio resources," but the methods of interacting in visual and oral media are still limiting (Thorne, 2008, p. 8). Expanding existent technologies to include features that offer additional visual support (through video of interlocutors) and written textual support (through simultaneous access to online dictionaries, pronunciation guides or translators) may make the devices more appealing to instructors and learners for use in synchronous communication.

The rapid increase in the popularity of mobile technologies and their noted potential for multidirectional communication makes investigation of how the technologies can be used for synchronous communication a top priority for second language researchers. While Colpaert (2004) forecasts that "the mobile hype will burst out as soon as tools become available allowing teachers and researchers to develop their own mobile applications and tools," passively awaiting the peak of the mobile revolution or the introduction of a ground-breaking technology is unacceptable (p. 262). As Robin predicts (2007), "the frontier in language learning and technology will not be found in what program does what better, but rather which students use the off-the-shelf technology to best facilitate their own learning in their own learning style" (p. 109). Teachers, researchers and curriculum designers must be proactive in seeking effective, practical methods that encourage the development of L2 oral proficiency in second language learners while transforming the students' learning experience into a seamless part of their daily lives.

ABOUT THE AUTHOR

Sarah Huffman is a first-year student in the doctoral program in Applied Linguistics and Technology at Iowa State University. She received an MA in Teaching English as a Second Language and Applied Linguistics from Iowa State. She has taught English as a Foreign Language in Slovakia and Mexico and English as a Second Language at a university in Virginia and Iowa State University. Sarah is currently a research assistant for the Research Writing Tutor project at Iowa State and is an editor for the Technology for Second Language Learning 2010 conference proceedings.

REFERENCES

Bachman, L. F., & Palmer, A.S. (2000). *Language testing in practice: Designing and developing useful language tests*. Oxford Applied Linguistics. Oxford: Oxford University Press.

Beatty, K. (2003). *Teaching and researching computer-assisted language learning*. Essex, England:Pearson Education Limited.

Chinnery, G. (2006). Emerging technologies: Going to the MALL, Mobile Assisted Language Learning. *Language Learning & Technology, 10*(1), 9-16.

Colpaert, J. (2004). From courseware to courseware? *Computer Assisted Language Learning, 17*(3), 261-266.

Hampel, R. & Hauck, M. (2006). Computer-mediated language learning: Making meaning in multimodal virtual learning spaces. *JALT CALL Journal, 2*(2), 3-18.

Harley, B. (1993). Instructional strategies and SLA in early French immersion. *Studies in Second Language Acquisition, 15*, 245-260.

Heins, B., Duensing, A., Stickler, U., & Batstone, C. (2007). Spoken interaction in online and face-to-face language tutorials. *Computer Assisted Language Learning, 20*(3), 279-295.

Kern, R. (1995). Restructuring classroom interaction with networked computers: Effects on quality and characteristics of language production. *Modern Language Journal, 79*(4), 457-476.

Kiernan, P.J., & Aizawa, K. (2004). Cell phones in task based learning: Are cell phones useful language learning tools? *ReCALL, 16*(1), 71-84.

Lankshear, C., Gee, J. P., Knobel, M., & Searle, C. (1997). *Changing literacies.* Philadelphia: Open University Press.

Linnell, J. (1995). Can negotiation provide a context for learning syntax in a second language? *Working Papers in Educational Linguistics, 11*(2), 83-103.

Murphy, L. (2008). Supporting learning autonomy: Developing practice through the production of courses for distance learners of French, German and Spanish. *Language Teaching Research, 12*(1), 83-102.

Norbrook, H., & Scott, P. (2003). Motivation in mobile modern foreign language learning. In J. Attewell, G. Da Bormida, M. Sharples, & C. Savill-Smith (Eds.), *MLEARN 2003: Learning with mobile devices* (pp.50-51). London: Learning and Skills Development Agency.

Payne, J. S. & Whitney, P. J. (2002). Developing L2 oral proficiency through synchronous CMC: Output, working memory, and interlanguage development. *CALICO Journal, 20*(1), 7-32.

Robin, R. (2007). Commentary: Learner-based listening and technological authenticity. *Language Learning & Technology 11(*1), 109-115.

Salaberry, M. R. (2001). The use of technology for second language learning and teaching: A retrospective. *The Modern Language Journal, 85*(i), 39-56.

Thorne, S. (2008). Computer-mediated communication. In N. Van Deusen-Schol & N.H.Hornberger (Eds.), *Encyclopedia of Language and Education, 2nd Edition, Vol. 4: Second and Foreign Language Education,* (pp. 1-12). Springer Science and Business Media LLC.

Thorton, P., & Houser, C. (2003). Using mobile web and video phones in English language teaching: Projects with Japanese college students. In B. Morrison, C. Green, & G. Motteram (Eds.), *Directions in CALL: Experience, experiments & evaluation* (pp. 207-224). Hong Kong: English Language Centre, Hong Kong Polytechnic University.

Van Dijk, J. (2005). *The deepening divide: Inequality in the information society.* London: Sage.

Volle, L. M. (2005). Analyzing oral skills in voice email and online interviews. *Language Learning & Technology, 9*(3), 146-163.

Wang, Y. (2004). Distance language learning: Interactivity and fourth-generation Internet-based videoconferencing. *CALICO Journal, 21*(2), 373-395.

Wang, Y. & Sun, C. (2001). Internet-based real time language education: Towards a fourth-generation distance education. *CALICO Journal, 18*(3), 539-561.

Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13*(2), 7-26.

Whitney, L. (2010). Cell phone subscriptions to hit 5 billion globally. *Mobile World Congress.* Retrieved April 27, 2010 from http://reviews.cnet.com/8301-13970_7-10454065-78.html.

Koffi, E. (2011). The pronunciation of /s/ in complex onset and coda clusters in Somali-accented English. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (pp. 130-143), Ames, IA: Iowa State University.

THE PRONUNCIATION OF /s/[ii] IN COMPLEX ONSET AND CODA CLUSTERS IN SOMALI-ACCENTED ENGLISH

Ettien Koffi, Saint Cloud State University

The pronunciation of /s/ in complex onset and coda clusters presents a formidable challenge to Somali English Language Learners (ELLs). There are two primary reasons why /s/ contributes to accentedness. First, in English /s/ occurs in a wide array of phonological environments in which it is not permissible in Somali. Secondly, many of the distributional patterns of the English /s/ violate the Sonority Sequencing Principle (SSP) and the Coda Condition, two widely attested constraints in world languages. This study shows that the Somali pronunciation of /s/ becomes less accented the more the speaker learns to violate these two constraints in the same way that native/proficient speakers of English do. Therefore, teachers can help their students produce /s/ successfully if they teach them two simple strategies: an exaggerated elongation of /s/ in complex onset and coda clusters and an exaggerated voicing of <-s> when it is added as a suffix to a voiced segment.

INTRODUCTION

The pronunciation of /s/ in complex onset and coda clusters presents a formidable challenge to Somali English Language Learners (ELLs). Complex onsets are understood in this paper as referring to syllables that contain two or more consecutive consonants before the nucleus. Similarly, complex codas are those that contain a sequence of two or more consonants after the nucleus. Faircloth and Faircloth (1973, p. 57) list /s/ as the fourth most frequent consonant in English, right after /n/, /t/ and /d/. It also occurs more frequently in onset and coda clusters than any other sound. In such environments, its pronunciations vary from [s] to [z] to [əz]. Whitley (2004) also notes that /s/ occurs in 87% of the world's languages. The frequency of /s/ in English and its pervasiveness in world languages can be both a source of positive and negative transfer. The segment /s/ is likely to be transferred positively because it has similar phonetic characteristics across a vast array of languages. However, /s/ can also be subject to interference because languages have various phonotactic constraints that govern its distribution. As a result, it is rare to find two languages where /s/ is pronounced identically in complex onset and coda clusters. This is especially true for the three Somali speakers whose pronunciation is the subject of this study. Because of the phonotactic differences between Somali and English, more often than not, the speakers fail to apply the rules that convert /s/ into [z] in the coda. In other instances, they epenthesize a schwa between complex onset and coda clusters involving /s/. Occasionally, they delete /s/ from complex clusters altogether. Not correctly applying the various rules that govern the pronunciations of the English /s/ often brings about an unmistakable

indexical feature of accentedness to Somali English. *The Sonority Sequencing Principle* and *The Coda Condition* are used in this paper to provide a principled account of this aspect of Somali pronunciation. These phonological principles will be explained later in the paper. The insights gathered from this analysis may lead to pedagogical strategies that teachers can use to help their Somali students improve their production of /s/ in various environments.

COMPARISON BETWEEN ENGLISH AND SOMALI SYLLABLE STRUCTURES

All languages organize phonemic materials into higher units called syllables. The basic syllable structure that has been attested in languages worldwide is the combination of a single consonant and a single vowel into a canonical structure represented as CV. However, languages are free to add more consonants before or after the vowel. For the purposes of comparison, let's contrast how English and Somali add consonants and vowels to the universal CV structure:

Table 1. *Comparison of Syllable Structures of English and Somali*

English Syllables	**Percentage**	**Somali Syllables**	**Percentage**
CV	32.49	CV	Not available
CVC	30.22	CVC[iii]	Not available
VC	16.34	VC	Not available
V	8.11	V	Not available
Total of simple onsets or codas	87.07		
CVCC	5.55		
CCVC	2.84		
CCV	2.64		
VCC	0.72		
CCVCC	0.60		
CCCVC	0.24		
CCCVCC	0.19		
CVCCC	0.12		
CCVCCC	0.02		
CCCV	0.01		
Total of complex onsets and codas	12.93		

The English data is taken from Faircloth and Faircloth (1973, p. 78) and the information about the canonical syllable structure of Somali comes from Saeed (1999, p. 25). It is worth noting right away that simple codas can become complex when the suffix <-s> is added to English words to form plural nouns, possessive noun phrases, or when the verb agrees with its third person singular subject in the present tense. So, if morphophonological rules are taken into account, the number of complex codas in English is significantly higher than those reported in

Table 1. Compared to English, Somali syllable structures are rather simple and straightforward: no complex onsets and no complex codas.

DATA ANALYSIS AND METHODOLOGY

The data that serves as the basis for this analysis comes from the George Mason University (GMU) Speech Accent Archive (http://accent.gmu.edu/browse_language.php). As of August 19, 2010, researchers at GMU have compiled 1,309 recorded readings of the text below:

> Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

The goal of the project (as stated on the GMU website) is the following:

> The speech accent archive uniformly presents a large set of speech samples from a variety of language backgrounds. Native and non-native speakers of English read the same paragraph and are carefully transcribed. The archive is used by people who wish to compare and analyze the accents of different English speakers.

Many of the recorded texts have been transcribed by trained graduate students using the International Phonetic Alphabet (IPA) characters and symbols but many more are awaiting transcription. There are six recordings by Somali speakers and three of them have been transcribed phonetically. It is these transcribed texts that serve as the data for my analysis of Somali-accented English.

Table 2. *IPA Transcriptions of Somali Data*

Speaker 1 female, Mogadishu (Speech Accent Archive)	**Speaker 2** male, Erigavo (Speech Accent Archive)	**Speaker 3** female, Borama (Speech Accent Archive)
[plis kal əstɛla æskə her tũ bi̤ĩŋk d̪oz̪ i̤ĩŋs wɪt̪ her fɹɔm ðə stɒr sikəs spũns ʌf fɹɛʃ snoʊ pis: faɪf s̪ stɪk θlæbəs ʌf blu ʧiz̪ ænd mɛbi ə̃ snæk for her bɹʌðer bab wi ɑlso nid ə̃ smɑl plæstik əsnæg ænd big˺ tʰɔɪ f fɔɹk for ðə kʰɪts ʃi kæ̃n əskjup d̪oz ĩŋk ɪn ðə θri ɹɛd bæg̪s æn wi wɪl go mið her wɛnəsdeɪ æt˺ tren steʃə n]	[pli:s kʰɔl əstɛla æskĩ her tu brĩŋ ðis θĩŋz wið her frõm dɪ stoɹ s̪ɪks ĩspũns of freʃ ĩsno pi:s faɪv t̪ik slæbz əv blu ʧi:s ẽn meɪbɪ e snæk for her braðɛr bob̪ wɪ ɔlso ni:d e smɔl plæstik snek ẽn eɪ bɪg tɔj frɔg for dɪ kidəs ʃɪ kæ̃n sku:p d̪i:z θɪŋz ĩntu θri rɛd bægs ẽn wɪ wɪl go mi:t her wĩnsd̪eɪ æt̪ə dɪ treĩn steʃə̃n]	[plis kɔl ɪstilɑ askə xɑr: tu briŋ d̪is θiŋks wɪθ xɑr: frʌm d̪i stɔɹ sɪkəs ɪspũns ŏf fərɛʃ: sno pis farɣ̊ θik ɪsɪslɪps ŏɣ̊ blu: ʧis æ̃n meɪbi ĕ sneɪk fɔr xɑr brɑðəɹ bop wĭ ɔlso nid ĕ smɔl blast̪ɪk snæk æ̃n ĕ bɪk tɔɪ fɹɔŋk fɔr ðə kɪrəs ʃi kæ̃n sku:p d̪is θiŋgəs ɪntũ t̪ri: rɛd̪˺ bægs æ̃n wi wɪl go mi:t xɑr wensdeɪ æt d̪ə treɪn steɪʃə̃n]

There are advantages and disadvantages in using a standardized text such as this one for phonological analysis. One main disadvantage is that the text may not be representative of the phenomenon under investigation. Though this is often true, in the present case, this disadvantage is mitigated by the fact that the text under consideration was carefully constructed to cover the main sounds of English that contribute to accentedness. In fact, all the phenomena under consideration are amply represented in the recording except for /s/ deletion in the coda. Moreover, the fact that the texts are transcribed by trained phoneticians is an added bonus because I do not have to rely on my own impressionistic judgments. According to the GMU website, utmost care has been used in transcribing all their texts. So, the transcribed texts are taken as prima facie evidence of Somali-accented English without second-guessing the accuracy of the data. I have listened to the recordings numerous times and have perceived only very minute and inconsequential disparities between the proposed IPA transcriptions and my own perception of the pronunciation of specific words. Furthermore, the disparities in question have to do with the transcription of vowels, not that of consonants. In addition to the information available of the GMU website, I have also consulted two M.A. theses (Lindsey, 2006; Admuson-Cisse, 2009) devoted to the pronunciation patterns of Somali ELLs. Now that all the caveats are firmly in place, let us embark on the analysis of the data.

The analysis of the recorded texts focuses only on the pronunciation of /s/ in complex onset and coda clusters. Though there are other phonological issues worthy of attention, they are not taken into consideration for this paper. The paper is organized into three main sections. The first deals with complex onset clusters, the second with complex coda clusters, and the third with suggested pedagogical strategies for addressing the pronunciation of /s/ in the environments mentioned in the two previous sections. The complex onset clusters represented in Table 3 will be the focus for our analysis in this section.

Table 3. *Complex Onset Clusters and Somali Realizations*

Words	**Clusters**	**Occurrences**	**Somali Realizations**
<Please>	[pl]	2	[plis]/[pli:s]/[plis]
<Stella>	[st]	3	[əstɛla]/[əstɛla]/[ɪstila]
<bring>	[br]	2	[bɹɪn̪k]/[brɪn̪]/[briŋ]
<from>	[fr]	3	[fɹɔm]/[frɒ̰m]/[frʌm]
<store>	[st]	3	[stɒr]/[stor]/[stɔɹ]
<spoons>	[sp]	1	[spu̪ns]/[ɪspu̪ns]/[ɪspu̪s]
<fresh>	[fr]	3	[fɹɛʃ]/[frɛʃ]/[fərɛʃ:]
<snow>	[sn]	3	[snoʊ]/[ɪsno]/[sno]
<slabs>	[sl]	1	[θlæbəs]/[slæbz]/[ɪsɪlɪps]

<blue>	[bl]	1	[blu]/[blu]/[blu:]
<snack>	[sn]	3	[snæk]/[snæk]/[sneɪk]
<brother>	[br]	2	[bɹʌðer]/[braðɛr]/[braðər]
<small>	[sm]	1	[smal/[smɔl]/[smɔl]
<plastic>	[pl]	2	[plæstik]/[plæstik]/[blastɪk]
<snake>	[sn]	3	[əsnæg]/[snek]/[snæk]
<frog>	[fr]	3	[fɔɹk]/[frɔg]/[fɪɔŋk]
<scoop>	[sk]	1	[əskjup]/[sku:p]/[sku:p]
<three>	[θr]	1	[θri]/[θri]/[tri:]
<train>	[tɹ]	1	[tren]/[trɐ̯ɪn]/[treɪn]
<station>	[st]	3	[steʃən]/[steʃə̯n]/[steɪʃən]

The phonetic transcriptions of words in the fifth column reflect the order in which Speakers 1, 2, and 3 said them. Though as many as 20 onset (initial) clusters appear in the data, in this paper we will concern ourselves only with those that violate a quasi-universal phonotactic constraint called the Sonority Sequencing Principle (SSP). The English onset clusters that violate the SSP are /sk/, /sp/, and /st/.

OVERVIEW OF THE SONORITY SEQUENCING PRINCIPLE (SSP)

The SSP is widely used by phonologists in theoretical studies of the syllable. It is stated in Roca and Johnson (1999, p. 266) as follows:

> The Sonority Sequencing Principle[iv] - The sonority profile of the syllable must rise until it peaks, and then it falls.

The SSP has attained the status of a universal principle even though it was not initially formulated as a language universal. Goldsmith (1990, p. 111) cautions that the SSP is only intended "as a necessary condition for basic syllabification, not a universal statement of possible syllables in any language." In spite of this cautionary note, most analysts today view the SSP as a universal principle or as an "emergent" universal constraint. The SSP owes its "elevated" status to the fact that it is applicable to a very large number of languages. Linguists have depended on sonority levels between sequences of sounds to describe specific languages and offer generalizations. Each phoneme is deemed to have a sonority scale, as explained by Ladefoged (2006, pp. 239-40): "The sonority of a sound is its loudness relative to that of other sounds with the same length, stress, and pitch." Taking cues from other linguists, Goldsmith (1990, p. 112) has assigned a fairly comprehensive system of numerical values to almost all

classes of phonemes. He defends his approach by arguing that "while there is considerable skepticism that the ultimate account of sonority is one based on an arithmetic system of this sort, there may be something right about an account that is sufficiently oriented to measuring sonority differences to be able to state unambiguously that liquids are halfway between obstruents and vowels." The analysis in this paper is based on Goldsmith's numerical values because they offer a quantifiable means by which one can account for Somali-accented pronunciation of /s/ in various complex clusters.

Table 4. *The Arithmetic of Sonority*[v]

Sounds	Sonority Indexes	Features
[a, æ, ə]	10	low vowel
[e, o]	9	mid vowels
[i, u, j, w]	8	high vowels
[r]	7	rhotic
[l]	6	liquid
[m, n, ŋ]	5	nasals
[s]	4	sibilant
[v, z, ð]	3	voiced fricatives
[f, θ]	2	voiceless fricatives
[tʃ, dʒ]	1.5	affricates[vi]
[b, d, g]	1	voiced stops
[p, t, k]	0.5	voiceless stops

English complex onset clusters behave in two ways with respect to the SSP. In most cases, they conform to it. In other cases, they violate it. Various attempts have been made to explain away the violation of the SSP. Roca and Johnson (1999, pp. 290, 488) resort to a suspicious historical argument to claim the /s/ in /sk/, /sp/, and /st/ onset clusters should be considered extrasyllabic, namely that it is a syllable all by itself. However, doing so violates another important syllabification principle, Exhaustive Syllabification, which requires that all non-sonorant segments be part of a syllable. Since /s/ is not a sonorant, it would be strange to make an exception for it and postulate that it is a syllable all by itself. For other attempts to deal with the violation of the SSP, see Cho and King (2003, p. 185).

The syllable structure rules of many world languages do not permit two or more consonants to occur in the onset of a syllable. In fact, English is in the company of only a small number of languages that allow two or three consonants at the beginning of syllables. Therefore, English complex onset clusters violate the SSP. This paper seeks to provide some insight on accented English by attempting to answer the following question: What happens when a person whose first language conforms to the SSP wants to learn a language such as English that violate it? Research done by Guffey (2002) and others on Spanish speakers learning English has found that the SSP is a major contributor to accentedness. The disparity between English and Spanish

complex onset clusters is the main reason why English words such as [spɔrt] <sport>, [stʌdi] <study>, and [ski] <ski> are pronounced by many Hispanics as [ɛspɔrt], [ɛstʌdi] and [ɛski] respectively. Hispanics who are learning English as a second language transfer their simpler onset structure into English. How do Somali speakers fare when they are confronted with English /sk/, /sp/, and /st/ onset clusters?

The Violation of the SSP and Somali Speakers

The data in Table 3 shed some light on the question raised in the previous paragraph. There are 30 complex onset clusters in the data containing an initial /s/. The three Somali speakers had a hard time pronouncing 10 of them accurately. This represents an error rate of 33.33%. More specifically, 15 of these 30 words begin with /sk/, /sp/ and /st/ clusters. The Somali speakers mispronounced six of them, that is, an error rate of 40%. A case in point is the pronunciation of [stɛlə] <Stella>. All three Somali speakers realized it either as [əstɛla], [əstela], or [ɪstila]. Let's see how the syllabification displays depict the sonority profile of <Stella> as produced by Minnesotan Speaker 143[vii] and that of Somali Speaker 1:

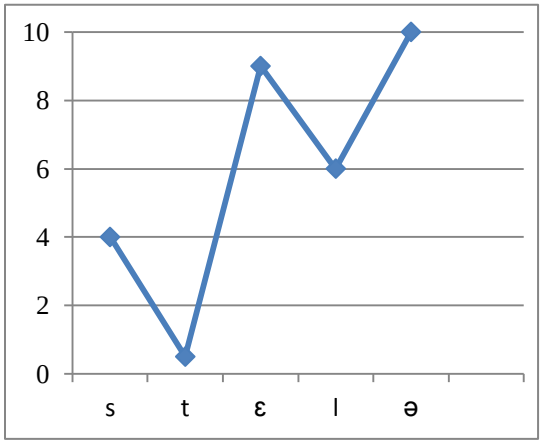


Figure 2. Syllabification display of [stɛlə] by speaker 143

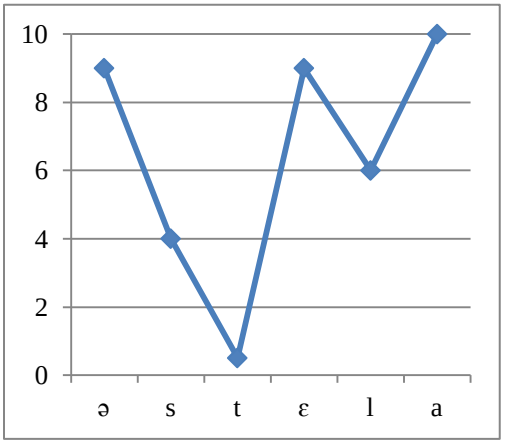


Figure 3. Syllabification display of [əstɛla] by Somali speaker 1

The pronunciation of Speaker 143 violates the SSP because the sonority index drops from [s] (4) to [t] (0.5). The SSP constraint requires the sonority profile to rise, not to drop. How do Somali speakers deal with this violation of the SSP? They completely resyllabify the word [stɛlə]. Instead of it being a two-syllable word [stɛ•lə], it becomes a three-syllable word in Somali-accented English [əs•tɛ•la]. The complex onset cluster [st] becomes [əs]. The restructured syllables in Somali English conform to the SSP. The insertion of the schwa allows [əs] to become a full-fledged syllable. The sound sequence [tɛ] also becomes its own syllable, and so is [la]. As noted earlier, the resyllabification of English words such as <scoop>, <spoon>, and

<Stella> that contain /sk/, /sp/, and /st/ in the onset accounts for 40% of complex onset errors. In general, if an English complex onset does not violate the SSP, Somali speakers have less difficulty with it. Thus, the pronunciation of /s/ in words such <snow>, <snack>, and <small> has a higher degree of accuracy. There are however, some surprising findings. Though, one would expect the speakers to have some difficulty with <store> and <station>, they did not. Conversely, even though the /s/ in <slab> does not violate the SSP, two speakers mispronounced the word. These "anomalies" show that Ladefoged (2006, p. 240) was right in saying that "a sonority theory of the syllable will not, however, account for all observed facts." Yet, the SSP can also help answer teachers' questions as to why Somalis generally insert the vowel [ə] at the beginning of /sk/, /sp/, and /st/ clusters.

THE PRONUNCIATION OF <-S> IN THE CODA

The three speakers produced a total of 30 coda clusters involving <-s>, as shown in Table 5. The pronunciation of <-s> (which may be either [s] or [z]) in 26 of its 30 occurrences shows a deviation from expected norms. This represents a mispronunciation rate of 86.6%. This underscores the fact that <-s> is more often mispronounced in coda clusters than in the onset clusters.

Table 5. *Complex Coda Clusters*

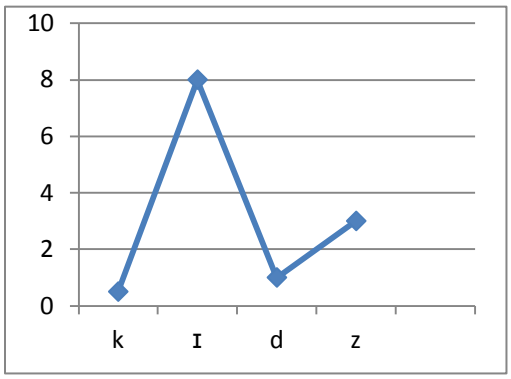
Words	Clusters	Occurrences	Somali Realization
<ask>	[sk]	1	[æskə]/[æskɪ]/[æskə]
<bring>	[ŋ]	1	[bɪi̯ŋk]/[brɪ̯ŋ]/[briŋ]
<things>	[ŋz]	2	[fi̯ŋs]/[θɪ̯ŋz]/[θɪŋks]
<six>	[ks]	1	[sikəs]/[sɪks]/[sɪkəs]
<spoons>	[nz]	1	[spu̯ns]/[ɪspu̯ns]/[ɪspu̯s]
<slabs>	[bz]	1	[θlæbəs]/[slæbz]/[ɪsɪlɪps]
<kids>	[dz]	1	[k^{h}ɪts]/[kidəs]/[kɪɾəs]
<things>	[ŋz]	2	[fi̯ŋk]/[θɪŋz]/[θiŋgəs]
<bags>	[gz]	2	[bægs]/[bægs]/[bægs]
<Wednesday>	[nz]	2	[wɛnəsdeɪ/wɪnsdeɪ/[wensdeɪ]

The pronunciation difficulties encountered by Somali speakers can best be accounted for by the Coda Condition. The SSP and the Coda Condition are like the two sides of the same coin. The SSP accounts for accented pronunciation in complex onsets while the Coda Condition sheds light on the pronunciation difficulties in complex codas. Kenstowicz (1994, p. 254) states the Coda Condition as follows:

Coda Condition - Codas fall in sonority from the nucleus.

Yavaș (2006, p. 139) elaborates on the Coda Condition by observing that "this means that optimal codas should have the sonority dropping as we move from C_1 to C_2." English can allow up to four consonants in the coda if the inflectional suffix <-s> is included in the count. According to the Coda Condition, the sonority must fall from the nucleus down to the closest consonant. The sonority must continue falling from that consonant until we reach the last consonant in a complex coda. If the sonority does not fall, then the Coda Condition has been violated. The coda structure of English syllable structures differs in significant respect from Somali coda structure, where only one consonant is allowed. As a result, pronunciation difficulties in the coda are to be expected.

The singular focus of this section of the paper is on the English plural suffix <-s> when it is added to syllable codas ending in /p, b, t, d, k, g/. The inflectional morpheme <-s> has three allophones: [s], [z], and [əz]. Only the first two pronunciations of <-s> violate the Coda Condition. The pronunciation [əz] does not violate the Coda Condition because [ə] has a sonority index 10 and [z] is 3, a drop of sonority. How do the three Somali speakers deal with the plural suffix <-s> when it is added to syllables whose codas are /p, b, t, d, k, g/? To illustrate this, let's compare the pronunciation of [kɪdz] <kids> by English Speaker 143 and by Somali Speaker 2.



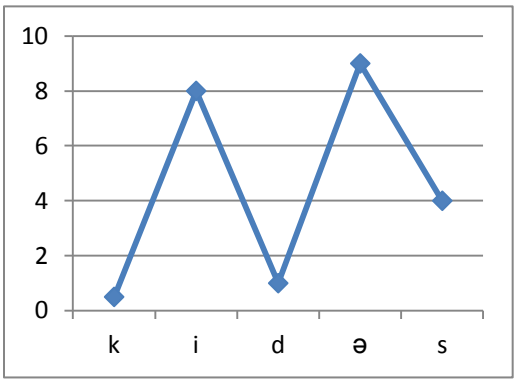


Figure 5. Syllabification display of [kɪdz] by Speaker 143

Figure 6. Syllabification display of [kidəs] by Somali speaker 2

The Coda Condition is violated in English because the sonority index drops from the nucleus [ɪ] (8) to [d] (1) and then rises again to [z] (3). How do Somali speakers generally respond when English words violate the Coda Condition? In the case of [kɪdz], they resyllabify the word by introducing a [ə] between [d] and [z]. The coda clusters in the data which violate the Coda Condition are [bz], [dz], and [gz] in the words [slæbz] <slabs>, [kɪdz] <kids>, and [bægz] <bags>. There are nine occurrences of such codas clusters in the data and eight of them are mispronounced. This represents an error rate of 88%. Moreover, in such cases, more often than not, Somali speakers fail to voice <-s> even when it occurs after a voiced consonant.

Pronunciation difficulties when the Coda Condition is violated extend beyond the inflection suffix <-s>. It also happens when the grapheme <x> occurs in the coda, as illustrated by the pronunciation of [sɪks] <six>:

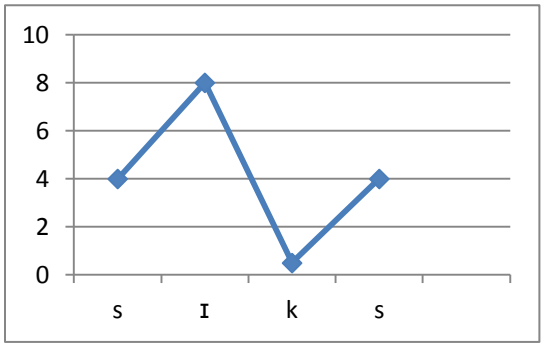


Figure 7. Syllabification display of [sɪks] by Speaker 143

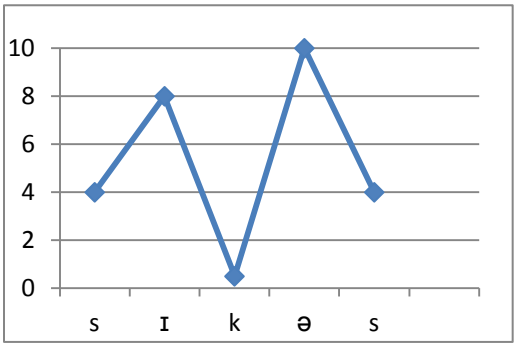


Figure 8. Syllabification display of [sɪkəs] by Somali speaker 3

The grapheme <x> consists of a sequence of two sounds [k] and [s]. In [sɪks]<six> the sonority level drops from [ɪ] (8) to [k] (0.5), and then rises again to [s] (4). The rise from [k] to [s] in the pronunciation of Speaker 143 violates the Coda Condition. Somali speakers 1 and 3 respond to this violation by resyllabifying [sɪks] into [si•kəs], as shown in Figure 8. The propensity of Somali speakers to resyllabify English coda clusters that violate the Coda Condition has also been attested in the spelling of college students. Admundson-Cissé (2009, p. 88) cites erroneous spellings of <parents> and <symbolism> as <parant**e**s> and <symbolis**i**m>, respectively.

The same resyllabification strategy is used in the pronunciation of the word [æsk] <ask> by all three speakers. They pronounce it as [æs•kə] or [æs•kɪ], that is, they insert an epenthetic vowel to break up the complex coda cluster [sk] even though this cluster does not violate the Coda Condition. The sonority value of /æ/ is 10, that of /s/ is 4, and the one for /k/ is 0.5. The sonority profile of [æsk] falls from the nucleus, as expected. Why then do the three Somali speakers resyllabify [æsk] as [æs•kə] or [æs•kɪ]? The answer is that Somalis tend to resyllabify coda clusters whether they violate the Coda Condition or not. Lindsey (2006, p. 55) also reports that vowel epenthesis is a very noticeable feature of Somali-accented English:

> In a recent conversation with ELL teachers, I was asked by several of them why Somali speakers of English add an /i/ to the end of their words. I asked for some examples and was given the following words: [fɪrsti] <first>, [hɪrti] for <hurt>, [ʃarki] for <shark>, [kaʊnti] <count>, and [læsti] for <last>. It is also interesting to note that these words also fit the pattern of errors in words that I have observed my students say.

None of these examples violates the Coda Condition, and yet Somali speakers insert an epenthetic /i/ to break up coda clusters.

PEDAGOGICAL IMPLICATIONS AND APPLICATIONS

The various rates of accented pronunciations of /s/ in complex onset clusters and <-s> in coda clusters are as follows:

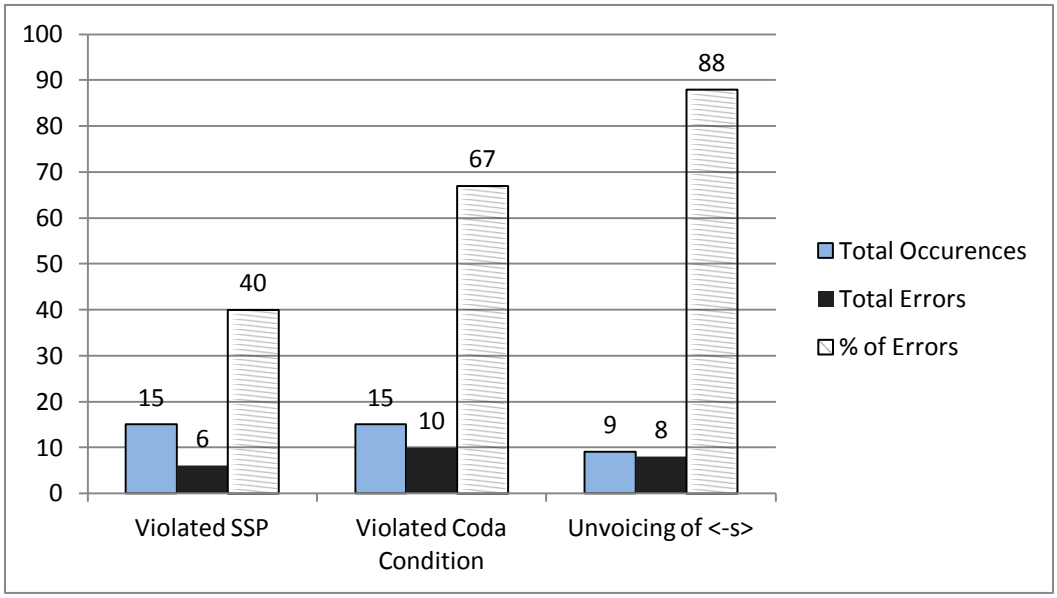


Figure 9. Ratio of errors per occurrence

The graphs show that coda cluster errors are more pervasive than onset cluster errors. Moreover, the percentage of errors increases tremendously when the suffix <-s> is not voiced. Normally, the morphophonological rule dictates that the suffix <-s> be pronounced [z] when it is added to voiced segment, be it a consonant or a vowel. Koffi (2010) and Wardhaugh (2010, p. 148) discuss the fact that failure to comply with morphophonological rules invites a negative evaluation of the speaker. Consequently, teachers of Somali ELLs should see to it that the <-s> that occurs in the coda after voiced segments is properly voiced. I suggest, following Celce-Murcia et al. (2010, pp. 104-5) that teachers exaggerate the pronunciation of <-s> after voiced segments as [z]. So, the <-s> in <peas> and <bags> should be taught to Somali students as [pi:zzzz] and [bægzzzz] respectively.

The pronunciation of /s/ in /sk/, /sp/ and /st/ in onset clusters also needs attention because such sequences violate the SSP. Again, following Celce-Murcia et al. (2010, pp. 104-5), I suggest that teachers teach their Somali students to exaggerate the initial /s/ in the clusters mentioned above. I propose that the initial /s/ in the words <scoop>, <spoon> and <Stella> be elongated as /ssssskup/, /sssspun/ and /sssstɛlə/. Pronouncing the initial /s/ accurately in these onset clusters is necessary lest Somali speakers are misunderstood by less cosmopolitan interlocutors who are not familiar with accented English. Word recognition experts who subscribe to the Cohort Model contend that the beginning of words is crucial for accurate word recognition. Byrd and Mintz (2010, p. 170) describe the Cohort Model as follows, “According to this theory, the first

phonological unit of a word, as recovered from the speech signal activates all the words in the listener's mental lexicon that begin with that sound." Since the Cohort Model relies on bottom-up processing, it can explain why, if an English word begins with an /sk/, /sp/, or /st/ cluster and if a Somali speaker pronounces it with an initial /ə/, this may cause the hearer to search in his/her mental lexicon for all the words that begin with a schwa. This miscue may actually cause some 250 milliseconds delay in recognizing the target word that the speaker is trying to say (Byrd & Mintz, 2010, p. 163). This may also cause the Somali speaker to have to repeat himself or herself more than once before being understood.

Coda cluster simplification by the deletion of one of the segments has been attested in Somali-accented English. Unfortunately, the text provided by the Speech Accent Archive does not give any evidence of it. Lindsey (2006, pp. 54-5), however, does not seem concerned by this pronunciation because "consonant cluster simplification by deletion strategies[viii] are used primarily by beginning students. As students progress to the intermediate level (where most of my students are), they tend to not simplify their consonant clusters by deletion. This has been confirmed by other ELL teachers of Somali who rarely speak of consonant cluster deletion as a serious problem." It should also be noted here that native and proficient speakers of American English do not think that coda cluster simplification by deletion is a serious problem because they too engage in it, as indicated by Yavaş (2006, p. 141):

> It would be appropriate to point out some modifications that are commonly observed with respect to deletions in the final clusters. When the word ending in a cluster is followed by a word that begins with a consonant, the final member of the cluster is deleted.

However, both Yavaş and Celce-Murcia et al. (2010, p. 107) add a very important caveat that needs to be heeded by teachers of Somali ELLs, namely that the suffix <-s> is rarely ever deleted by native/proficient speakers of English because of the vital grammatical load that it carries.

SUMMARY

The Sonority Sequencing Principle and the Coda Condition lend support to the notion that aspects of Universal Grammar may be responsible for accentedness. Somali speakers' pronunciation of complex onset clusters that begin with /s/ are less problematic when such words conform to the SSP. Similarly, complex coda clusters that conform to the Coda Condition are less difficult to pronounce than those that do not. If a teacher has Somali students in his/her class, all he/she needs to do is teach the students to explicitly and consistently violate the quasi universal SSP and Coda Condition constraints. The pedagogical strategy that seems to work best in this situation consists in teaching learners to elongate and exaggerate /s/ in /sk/, /sp/ and /st/ clusters into [sssssk], [ssssp], and [ssssst]. As for <-s> in coda clusters, it should not only be elongated, but it should also be voiced. So, the words [bægz] and [pi:z] would be pronounced respectively as [bægzzzz] and [pi:zzzz].

[ii] The following conventions are used throughout the paper: < > represents the orthographic transcription of a grapheme, / / stands for the phonemic transcription, and [ ] for the phonetic transcription. Their use in the paper is not arbitrary but corresponds to the type of argumentation being presented.

[iii] Saeed (1999, p. 17) provides the following minimal pair [war•an] (spear) and [warr•an] (tell news) to show that a CVCC pattern is possible. However, in all such cases, the two contiguous CCs represent a lengthened consonant which can be transcribed phonemically as /C:/. Consonant lengthening is phonemic in some East African languages, including Somali.

[iv] For CV languages, the sonority profile of the syllable rises and peaks but it does not fall. So, the SSP applies mostly to CVC languages. Since English and Somali have CVC syllables, the SSP applies to both languages.

[v] Various sonority values have been proposed. Roca and Johnson (1999, p. 288) propose a slightly different system from the one given here. Guffey (2002, p. 3, 5, 10) lists at least four other scales. The one proposed by Goldsmith is used in this analysis because it is the most comprehensive. I have augmented Goldsmith's list of phonemes by including information from Yavaş (2006, p. 131).

[vi] Goldsmith omits to assign numerical values to affricates on p. 112. However, on the previous page he ranks their sonority level between that of stops and fricatives. Consequently, assigning the value of 1.5 is fairly accurate. Guffey (2002, p. 10) makes a similar observation.

[vii] The Speech Accent Archive identifies a 42 year-old male speaker from St. Paul, Minnesota as speaker 143. His speech is used throughout this paper as model of speech that Somalis who live in the state are more likely to hear.

[viii] "In the coda" is intended by the author but was not explicitly stated.

ABOUT THE AUTHOR

Ettien Koffi teaches Interlingual Phonetics and Phonology, Pedagogical Grammar, and Sociolinguistics in the graduate TESL program at Saint Cloud State University. His areas of specialization and interest are in the application of phonological theory to the elaboration of orthographies for unwritten languages and the acoustic/phonological study of dialect variation.

REFERENCES

Admuson-Cisse, A. (2009). *The Effects of Pronunciation on Somali Spelling Errors*. MA Thesis. Saint Cloud State University: St. Cloud, MN.

Byrd, D. & Mintz, T. (2010). *Discovering Speech, Words, and Mind*. Malden, MA: Wiley-Blackwell.

Celce-Murcia, M., Brinton, D. & Goodwin, J. with Grice, B. (2000). *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages, 2nd Ed.*. New York, NY: Cambridge University Press.

Cho, Y. & King, T. (2003). Semisyllables and Universal Syllabification. In C. Féry & R. van de Vijver (Eds.) *The Syllable in Optimality Theory*, pp. 183-212. New York, NY: Cambridge University Press.

Faircloth, S. & Faircloth, M. (1973). *Phonetic Science: A Program of Instruction*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Cambridge, MA: Basil Blackwell, Inc.

Guffey, K. (2002). *Spanish Syllable Structure*. Laham, NY: University Press of America.

Kenstowicz, M. (1994).*Phonology in Generative Grammar*. Malden, MA: Blackwell Publishers.

Koffi, E. (2010). The Pronunciation of <-ED> in Coda Clusters in Somali-Accented English. In J. Levis & K. LeVelle (Eds.) *Proceedings from the 1st Conference of Pronunciation in Second Language Learning and Teaching*, pp. 119-134. Ames, IA: Iowa State University.

Ladefoged, P. (2006). *A Course in Phonetics*. Fifth Edition. New York, NY: Thomson-Wadsworth.

Lindsey, M. (2006). *An ELL Teacher's Guide to Somali-English Interlingual Phonology*. MA Thesis. Saint Cloud State University: St. Cloud, MN.

Roca, I. & Johnson, W (1999). *A Course in Phonology*. Malden, MA: Blackwell

Saeed, J. (1999). *Somali*. Philadelphia: John Benjamins Publishing Company.

Wardhaugh, R. (2010). *An Introduction to Sociolinguistics*, 5th Edition. Malden, MA: Wiley-Blackwell Publishers.

Whitley, S. M. (2004). *The Phonicon: A Multimedia Digital Encyclopedia of the Sounds of The World's Languages*. CD Rom. Dallas, Texas: SIL International.

Yavaș, M. (2006). *Applied English Phonology*. 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.

ACQUISITION OF L2 PHONOLOGY IN ADVANCED LEARNERS: DOES INSTRUCTION MAKE A DIFFERENCE?

Anita Saalfeld, University of Nebraska-Omaha

Research in second language phonology suggests that direct pronunciation instruction may improve students' pronunciation (Derwing, Munro & Wiebe, 1997; 1998; Lord, 2005). For example, Lord's (2005) study of the instructional effects of a Spanish phonetics course on advanced Spanish learners' acquisition of phonology showed that learners improved in their pronunciation over the course of a semester, but one of the study's limitations was that there was no control group. Therefore, it is unclear whether improvement was related to the phonetics course, or merely a result of being enrolled in any advanced Spanish course.

This study presents a preliminary analysis of one feature, Spanish stress, based on a replication of Lord's (2005) study, but includes a control group enrolled in advanced Spanish courses (n = 17), as well as the experimental group enrolled in a Spanish phonetics course (n = 11). All participants were enrolled in third and fourth year Spanish courses; they completed an initial recording during the first four weeks of the semester, and a final recording during the last four weeks of the semester. Results indicate that the students who most need pronunciation instruction are not enrolled in phonetics classes.

INTRODUCTION

As many researchers in L2 phonology have noted over the last fifteen years or so (Arteaga, 2000; Derwing & Munro, 2005; Elliott, 1997; González-Bueno, 2001; Hurtado & Estrada, 2010; Levis, 2005; Lord, 2005; Morin, 2007), pronunciation has been largely neglected in the L2 classroom. While there are undoubtedly a variety of reasons for this, Derwing & Munro (2005) note that the trend away from teaching pronunciation can probably be principally attributed to a shift away from audiolingualism, as well as a result of research emerging in the late 1960s and early 1970s suggesting that it was impossible for post-pubescent learners to acquire a native-like accent. Furthermore, communicative language teaching de-emphasizes drills of the kind that are usually used to practice pronunciation (repetition, for example).

However, this shift away from teaching pronunciation has left language learners at a disadvantage, because they are not being taught pronunciation in traditional language classrooms. As Levis (2005) and Derwing & Munro (2009) note, there is now an accent reduction industry, which promises to eliminate foreign accent, but there is no research in L2 pronunciation pedagogy to support the claims of the industry; the majority of the evidence instead indicates that attainment of a native accent in a second language is not a realistic goal for most language learners.

Nonetheless, this should not result in pronunciation instruction being excluded from the curriculum. The very existence of an industry promising (however fraudulently) to eliminate foreign accent indicates that learners want to improve their own pronunciation.[1] Moreover,

while not all aspects of pronunciation must be accurate in order for a speaker to be intelligible, there must certainly be some minimal degree of competence in pronunciation in order for communication to occur. For example, in a beginning Spanish classroom discussing weekend activities, a student was asked if he washed clothes on Saturdays. ("¿Lavas la ropa los sábados?") The student responded, befuddled, "Did you just ask me if I wore clothes on Saturdays?" (He was promptly reassured that the question was asking him if he *washed* clothes on Saturdays.) In Spanish, the verb "lavar" is "to wash," and the verb "llevar" is "to wear." The phonological similarity leads many students to mispronounce both words, to the point that they misinterpret basic questions in bizarre (though funny) ways.

By themselves, the communicative importance of pronunciation and learners' desires to master it is still not enough to justify its inclusion in foreign language instruction; we must also see evidence that teaching pronunciation is effective. Recent research investigating a number of features in various languages at different levels of instruction has shown that teaching pronunciation produces improvement.

In Spanish, Elliott (1997) and González-Bueno (1997) examined the effects of instruction on pronunciation in beginning and intermediate language courses. While participants did not demonstrate improvement in all of the features that were studied, they did demonstrate improvement in some of them. Also in Spanish, Lord (2005) investigated the effects of instruction on the pronunciation of students enrolled in a Spanish phonetics course. There were 17 learners enrolled in an upper-division Spanish phonetics course, as well as 10 native Spanish speakers that served as a control group. The study targeted various segmental features in Spanish, and the results indicated that although learners were still easily distinguishable from native speakers after a semester of instruction, their pronunciation improved noticeably over the course of the semester.

Lord (2008), Hahn (2004), Wang, Spence, Jongman, and Sereno (1999), and Derwing, Munro & Wiebe (1997, 1998) all examined instructional effects on acquisition of suprasegmental features in various languages (Spanish, English, and Mandarin), and participants in all studies demonstrated improvement after training.

The present study is a replication of Lord's (2005) study investigating the effects of instruction in a Spanish phonetics course. As noted above, results indicated that students improved their pronunciation over the course of the semester, but an important limitation of the study was the lack of a control group. Although it is extremely probable that the pronunciation instruction was the catalyst for improvement, without a control group, it is impossible to know for certain. With this in mind, the objective of the current study was to investigate whether improvement in pronunciation would occur in learners enrolled in a Spanish phonetics course, and if so, whether that improvement could be attributed to the phonetics course, or to being enrolled in any upper-level Spanish course. Features targeted for analysis in the full study include both segmental and suprasegmental features, but the current study focuses on lexical stress.

METHOD

Participants

Participants were 28 Spanish majors and minors at a large, public Midwestern university. The experimental group was made up of students enrolled in Spanish phonetics (n=11), and the control group (n=17) consisted of students enrolled in other upper-level Spanish courses. Most students in both groups were enrolled in at least two Spanish courses. Participants in the experimental group completed the tasks as part of their regular coursework; students in the control group made individual appointments with the researcher to make the recordings, and received a set of Spanish flashcards as compensation for their participation.

Design

The same text that was used in the Lord (2005) study was used for the present study (Appendix A). All participants completed pre-test and post-test recordings of the text. The variable under investigation for the current analysis was accuracy of stress placement. In order to judge stress placement, the researcher counted the number of words in the text with lexical stress containing two or more syllables (106 total), and totaled the number of words with incorrect stress placement for each participant on the basis of auditory discrimination.[2]

Procedure

Students enrolled in the Spanish phonetics course submitted a total of seven recorded homework assignments as part of their regular homework. The first and last recordings were the pre-test and post-test recordings of the experimental text. The researcher recruited students from other third and fourth-year Spanish courses to arrange appointments to record the text. All pre-test recordings were completed during the first three weeks of the semester, and all post-test recordings were completed during the final three weeks of the semester.

Over the course of the semester, students enrolled in the phonetics course received instruction on various aspects of Spanish pronunciation, including vowels, consonants, allophonic variation, stress, intonation, and dialectal variation. In addition to completing the seven recorded homework assignments, they took quizzes nearly every week on the content discussed during the previous week, and completed six written homework assignments, which included identifications, transcriptions, definitions, and more. Participants in the control group were enrolled in a variety of courses, such as literature, civilization, composition and conversation. None of the courses included regular, specific pronunciation instruction (though incidental pronunciation guidance may have occurred).

RESULTS AND DISCUSSION

Results from pre-test comparisons indicated that the experimental group was more accurate at producing correct word stress than the control group at the outset of the experiment (see Table 1). An independent samples t-test demonstrated that this difference was statistically significant ($t(26) = -2.23$, $p = .03$). Accuracy ratings for each group indicated that the experimental group produced correct word stress at a rate of 96.83% (102.64/106 words), and the control group produced correct word stress at a rate of 93.34% (98.94/106 words). While these figures are very high, they are interesting in light of the fact that Spanish stress is almost completely predictable based on orthography, and the words included in the experimental text did not present any exceptions. More interesting still is the fact that students enrolled in the phonetics

class already produced word stress with greater accuracy than the control group at the outset of the experiment. As is the case in many institutions, Spanish phonetics is an elective course, and is not required to graduate with a major or minor in Spanish. It is not a big surprise, then, that the students enrolled in the class demonstrated more pronunciation accuracy than those students who were not enrolled in the class; they selected the elective course out of an interest in improving their pronunciation, and this interest had already distinguished them from peers enrolled in other classes.

Table 1. *Pre-test means of accurately stressed words by group*

Group	N	Mean	SD	Accuracy rating
Experimental	11	102.64	3.26	96.83%
Control	17	98.94	4.80	93.34%

Possible total of 106 words.

Difference between groups was statistically significant ($t(26) = -2.23$, $p = .03$)

Post-test results revealed a similar pattern. Neither group demonstrated improvement in accuracy of stress production from pre-test to post-test, which was expected for the control group given the lack of instruction. However, considering the high accuracy rates of both groups, but especially of the experimental group, on the pre-test, it is probable that the lack of results is due to a ceiling effect. As Table 2 shows, the mean accuracy for the experimental group went up by .09, from 102.64 to 102.73, and the mean accuracy for the control group stayed the same at 98.94. A repeated measures ANOVA showed an effect for group ($F(1,26) = 5.33$, $p = .03$), but not for time ($F(1,26) = 0.01$, $p = .92$). Analyses of other features are still in progress, and will hopefully reveal improvement in the experimental group, especially in areas where there is more room for improvement. In particular, based on observations made while grading recorded homework assignments, participants will probably demonstrate improvement in vowel quality over the course the semester.

Table 2. *Pre-test and post-test means of accurately stressed words by group*

Group	N	Pre-test mean	Pre-test SD	Pre-test accuracy rating	Post-test mean	Post-test SD	Post-test accuracy rating
Experimental	11	102.64	3.26	96.83%	102.73	3.64	96.92%
Control	17	98.94	4.80	93.34%	98.94	4.76	93.34%

Possible total of 106 words.

Repeated-measures ANOVA indicated a difference between groups ($F(1,26) = 5.33$, $p = .03$), but no effect for time ($F(1,26) = 0.01$, $p = .92$)

Given the ceiling effect evident in the experimental group results, these results are not as discouraging as they otherwise would be. Rather, a key area of concern seems to be the self-selection bias that appeared in the experimental group. Specifically, the fact that students in the control group chose not to enroll in the Spanish phonetics course is a challenge for pronunciation instruction, since these students are the ones who need instruction the most. As in many programs, pronunciation is not emphasized in beginning Spanish language courses. While there are probably many reasons for this, one of the main reasons is almost certainly the relative orthographic transparency of Spanish; there are only five pure vowels, and most letters only correspond to one sound. Since there are many possible topics for instruction, the relative transparency of the Spanish alphabet means that instructors feel more comfortable omitting this information from their classroom instruction, and students who choose not to take Spanish phonetics may acquire a degree in Spanish without ever receiving pronunciation instruction.

The lack of pronunciation instruction throughout the curriculum means that students in upper-level Spanish courses, such as the phonetics course, arrive with entrenched errors in pronunciation. For example, although the participants enrolled in the phonetics class performed better than participants not enrolled in the phonetics class, there were a number of problems in the pronunciation of vowels. For example, a number of students reduced many unaccented vowels to [□], reduced /i/ to [⋂], /o/ to [˙·.] or [□], and frequently pronounced /u/ as [ju]. Given that there are only five vowels in Spanish, the degree to which the mis-pronunciation of vowels predominated in the phonetics course was astonishing. Because a substantial amount of time needed to be dedicated to correcting students' pronunciation of vowels, there was considerably less time to discuss and practice other features in the amount of detail originally planned.

Within the broader framework of SLA research, these results appear to corroborate Schmidt's (1990) Noticing Hypothesis, as well as work done by Leow (2000), which indicates that learners must be aware of a feature in order for it to become part of their linguistic system. Although the study was not designed to investigate attention and awareness, it appears that an awareness of the need to improve their pronunciation prompted students to enroll in the phonetics course. This awareness may be the reason for which students in the phonetics course displayed more accuracy in pronunciation at the outset than students not enrolled in the phonetics course. While there are indubitably many reasons why a student would choose not to enroll in a phonetics course, it is likely that a key reason is the belief that the course will not be beneficial; in other words, students are unaware of how far their pronunciation deviates from the norm and thus have no motivation to work on it.

The pedagogical implications of the results point to the need for pronunciation instruction integrated throughout the curriculum in second language instruction. At a minimum, in a language such as Spanish with only five vowels, students should not be systematically mispronouncing vowels at advanced levels of instruction. However, there is no easy solution to this problem. Derwing & Munro (2009) have noted that most language teachers do not learn how to teach pronunciation when they learn to teach other features in the language, and thus do

not feel comfortable or able to provide pronunciation instruction. Burgess & Spencer (2000) reported that even teachers enrolled in master's-level courses felt "ignorant of phonology, though they are aware of the need for helping learners with their pronunciation" (p. 207). Morin (2007) specifically notes the lack of preparation in teaching pronunciation for students planning to teach Spanish.

In order to address this issue, language education programs need to include pronunciation instruction in their curricula so that teachers are trained to work with students to improve pronunciation at all levels of language instruction. This would ideally prevent a number of pronunciation errors from becoming entrenched, and would allow an upper-level phonetics course to provide advanced, rather than remedial instruction to correct ingrained errors.

Another possible solution for language programs is to make a phonetics course (or another course focusing on pronunciation) a requirement in programs for foreign language majors and minors. While student evaluations indicated that most students did not care for the more theoretical parts of the course, all students felt that the pronunciation instruction given was valuable, and several students indicated that they had never received any type of pronunciation instruction in previous classes. One student commented: "I found this course to be more practical than some others that are required for Spanish majors. It helps with pronunciation and various problems of native [*sic*] speakers. In my opinion and experience, I have gained much more from it than literature classes." Another student noted that the class was very helpful, but that "I thought this was stuff I needed to know in my 2000 level [second-year] courses."

One possible solution for teachers wishing to offer pronunciation instruction but that feel unable to teach it is to give assignments that focus on pronunciation. For example, an instructor might make a recording of a native speaker reading a word list, series of sentences, or a short paragraph, and make the recording available to students. Students would then be required to listen to the recording and be directed to match target sounds as closely as possible. As technology continues to develop and become more affordable and accessible to non-specialists, it may also be possible to automate pronunciation training.

LIMITATIONS

As with any study, there are limitations that need to be addressed in future studies. One key limitation of this finding was that the number of Spanish classes that each student had taken and was taking at the time of the study was not controlled, and therefore, it is possible that the students enrolled in the phonetics course simply had the advantage of having completed more instructional hours in Spanish than students in the control group. For future studies, it would be advisable to collect background data about students' experience with Spanish in order to control for the amount of exposure to the language.

In addition, for future studies it may be worthwhile to collect survey data on attitudes and beliefs about acquiring L2 pronunciation. As an anonymous reviewer noted, it may be that students who chose not to enroll the phonetics course simply believed that it ultimately would not improve their pronunciation, and while research results in the broader SLA community indicate that instruction can improve pronunciation, the results of this particular study do not contradict that belief.

CONCLUSIONS

While recent research has indicated that pronunciation instruction is effective, it is still largely absent from language instruction at all levels. Because courses in phonetics for foreign language majors are often electives, and not required courses, many of the students who would benefit most from pronunciation instruction may choose not to take the course, and may graduate with a degree in a foreign language without being able to pronounce basic sounds correctly. To remedy this situation, as Derwing & Munro (2005) note, there is a need for more research on teaching pronunciation, as well as for more collaboration between researchers and teachers, so that research findings inform teaching practices.

NOTES

1. As an anonymous reviewer pointed out, the accent reduction industry is almost exclusively aimed at reducing a foreign accent in English. The lack of this industry in other languages is probably due to the fact that English is currently the international business language, and thus generates the largest amount of revenue. However, it may indicate that learners of other languages do not care if their speech is accented.

2. Analyses of other features are being conducted using the program Praat, as well as native speaker rater judgments, and are still in progress.

ABOUT THE AUTHOR

Anita Saalfeld is Assistant Professor of Spanish at the University of Nebraska-Omaha. She received her Ph.D. in Hispanic Linguistics from the University of Illinois at Urbana-Champaign. Her research interests are second language phonology and the use of technology in the language classroom.

Anita Saalfeld
University of Nebraska-Omaha
301C Arts & Sciences Hall
6001 Dodge St.
Omaha, NE 68182
402-554-3298
anitasaalfeld@unomaha.edu

REFERENCES

Arteaga, D. L. (2000). Articulatory phonetics in the first-year Spanish classroom. *Hispania*, *84*,339-354.

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

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

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. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48,* 393-410.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1997). Pronunciation instruction for "fossilized" learners: Can it help? *Applied Language Learning, 8,* 217-235.

Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania, 80,* 95-108.

González-Bueno, M. (2001). Pronunciation teaching component in SL/FL education programs: Training teachers to teach pronunciation. *Applied Language Learning, 12,* 133-146.

González-Bueno, M. (1997). The effects of formal instruction on the acquisition of Spanish stop consonants. In W. R. Glass & A. T. Pérez-Leroux (Eds.) *Contemporary Perspectives on the Acquisition of Spanish, Volume 2: Production, Processing and Comprehension* (pp. 57-75). Somerville, MA: Cascadilla Press.

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

Hurtado, L. M., & Estrada, C. (2010). Factors influencing the second language acquisition of Spanish vibrants. *Modern Language Journal, 94,* 74-86.

Leow, R. P. (2000). A study of the role of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition, 22,* 557-584.

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

Lord, G. (2008). Podcasting communities and second language pronunciation. *Foreign Language Annals, 41,* 364-379.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania, 88,* 257-267.

Morin, R. (2007). A neglected aspect of the standards: Preparing foreign language Spanish teachers to teach pronunciation. *Foreign Language Annals, 40,* 342-360.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11,* 129-158.

Wang, X., 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 A

Text read by participants

Poco a poco, a medida que la noche fue avanzando, la presencia de la fotografía empezó a hacerse más molesta y obsesiva cada vez. Concentré la mirada en la espiral del fuego. Cerré los ojos tratando de dormir. Pero todo era inútil. Los ojos amarillos de Sabina me miraban. Su soledad antigua se extendía como una mancha húmeda por toda la pared. Pronto entendí que la tranquilidad y el sueño de horas antes serían ya imposibles mientras aquel viejo retrato siguiera frente a mí. La perra despertó sobresaltada, y se quedó mirándome sin entender muy bien. Yo estaba ya junto al escaño, nervioso y aturdido, pero dispuesto a poner fin a aquella situación. El recuerdo cercano de la soga me empujaba. El temor a la locura y al insomnio había comenzado a apoderarse de mí. Cogí el retrato entre las manos y lo miré otra vez: Sabina sonreía con una gran tristeza, sus ojos me miraban como si aún pudieran ver. Y, en la desolación extrema de aquel andén vacío – vacío para siempre -, su soledad de entonces atravesó mi corazón. Sé que nadie jamás me creería, pero, mientras se consumía entre las llamas, su voz inconfundible me llamaba por mi nombre, sus ojos me miraban pidiéndome perdón.

From *La lluvia amarilla*, by Julio Llamazares, Quoted from Lord (2005)

Wallen, S. & Fox, R. A. (2011). Vowel spaces in bilingual Haitian American kindergartners. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 153-168), Ames, IA: Iowa State University.

VOWEL SPACES IN BILINGUAL HAITIAN AMERICAN KINDERGARTNERS

Stacey Wallen, Georgia State University

Robert A. Fox, The Ohio State University

With the number of English Language Learners (ELL) in the United States, it has become important for communication professionals to understand how speech and language skills develop in this population in order to correctly differentiate a "communication disorder" from a "communication difference." In an effort to provide information on young ELLs, this study provides an acoustic description of Kreyol and English vowels spoken by monolingual and bilingual Haitian American kindergartners. Ten kindergarteners of Haitian descent produced words containing Kreyol and English vowels in either CV, CVC or CVCV contexts. Their productions were compared to eight non-Haitian children from the same region. The frequencies of the first two formants were extracted at the vowels' midpoint and compared between three groups: Haitian American monolingual English speakers (HAM), Haitian American bilingual (English/Kreyol) speakers (HAB), and Non-Haitian speakers (NH). Results for Kreyol vowels provide a first-time acoustic description of the Kreyol vowel space. Results for English vowels reveal significant differences in the production of one vowel, /o/, between HAM and HAB speakers. No significant differences in the English vowel spaces of bilingual Haitian American children were observed when compared to vowel spaces of their non-Haitian native counterparts.

INTRODUCTION

Children of Haitian descent living in the United States are part of the population that Portes and Rumbout (2001) call "children of the second generation"--first generation Americans either born in the United States to immigrant parents or brought to this country before turning 18. These children have connections to both the American and Haitian culture and can communicate fluently in either English or Haitian Kreyol. This fluency in English can have an effect on how professionals, especially teachers and speech-language pathologists, in the United States view speech and language data in this population. In order to effectively diagnose and treat communication disorders in people from different groups, speech language pathologists must know about what's "typical" in the language groups that they serve. They must also understand how that process of development changes when applied to a bilingual learner.

The National Clearinghouse for English Language Acquisition and Language Instruction Educational Programs (NCELA) reports that Haitian Kreyol is one of the most common languages spoken by school-aged English language learners in the United States; it is ranked sixth in language backgrounds of English Language Learners (ELL) (Kindler, 2002). Because young school-aged children make up a large percentage of the children referred for articulation services, knowing how bilingual speakers (Kreyol/English speakers in this case) produce

phonemes is important in determining which child will need articulation services. Although speech-language pathologists are trained to take the child's native language into consideration when assessing speech and language skills, the lack of objective developmental data on bilingual speakers too often causes the speech-language specialist to treat children that are bilingual as monolingual speakers of their native or second language. Therapists either treat the bilingual child using the data available in their native language (not considering the fact that they know/speak English), or they treat the child as a native English speaker of the mainstream dialect (not taking into consideration the child's use of their native language). This can lead to possible over-diagnosis or under-diagnosis of a communication disorder.

When we investigate phonological development in adult second language learners, we find that the sounds in the L2 that are the most salient perceptually will be produced first (Gorman & Kester, 2003). This is applicable for different languages or different dialects. The Perceptual Assimilation Model (Best, 1995) provides an explanation of how and why this occurs. This model relates perceptual saliency to perceived vocal tract constrictions. A listener perceives the constrictions of the vocal tract that relate to their native language, judging similarities between the sound structures of L2 and L1. The listener then associates the non-native sound to the native sound that most closely resembles it.

Flege (1995) defined the "production" side of this phenomenon via the Speech Learning Model. The Speech Learning Model (SLM) also explains the changes that occur in L1 and L2 phonological systems when the two languages interact. It asserts that as L1 and L2 interact, performance in both languages is affected. Different factors, including age of acquisition, length of exposure, access to, and use of the L2 determine how the L1 and L2 will interact. Flege, MacKay and Meador (1999), examined age of arrival (AOA) and use of L1 as factors in native-like perception and production in bilingual (Italian/English) speakers. They determined that the earlier bilingual subjects began speaking L2 (English), the more likely it was that their L2 productions were like that of a native English speaker. Subjects learning English at a later age were more heavily influenced by their L1 (Italian) and had a tendency of producing English phonemes more consistent with phonemes in Italian. Continued use of the L1 was not found to be significant in native-like production or perception in this particular study. However, other studies have found that increased use of the L1 lead to more "accented" productions of the L2 (Flege, 1999).

Haitian Kreyol is a member of the "Atlantic French-based Creoles". Approximately 90% of its lexicon is derived from 16th-18th century French, but its morphology, semantics and syntax are more closely related to West African languages (Savain, 1999; St. Fort, 2000). Kreyol, the primary language spoken by about 90-95% of Haitians, was an oral language until 1980, when an official orthographic system was developed. Table 1 provides the 30 symbols that make up the Haitian Kreyol alphabet. Kreyol became an official language of Haiti in 1987; however, the use of Kreyol for official business and as the standard written medium has not been prevalent (Reagan, 2005; St. Fort, 2000). There are ten vowels in Kreyol's phonemic inventory (/i, ɛ, e, ɑ, u, o, ɔ, ɑ̃, õ, ɛ̃/), including three nasal vowels and seven non-nasal vowels (Savain, 1999; St. Fort, 2000; Muysken & Veenstra, 1995; Tinelli, 1981). Few studies written in English (Tinelli, 1981) have done an in-depth analysis of Kreyol phonology.

Table 1: *Kreyol Alphabet*

Consonants	Vowels	Semi-vowels
[b, ch, d, f, g, h, j, k, l, m, p, r, s, t, v, z]	[a, an, e, é, en, i, o, ó, on, ou]	[ui, w, y]

If we consider the fact that an older speaker learning a second language is less likely to sound like a native speaker of the L2, what happens when a younger speaker learns a second language simultaneously with her native language? Goldstein (2001) notes that the phonological system of a child that speaks Spanish can influence that child‘s production of English. However, to date, there are no known studies that look at this possibility for Haitian American children that are fluent in or influenced by English and Haitian Kreyol.

To provide developmental information on ELLs , the goal of this study is to provide a description of the speech patterns of Haitian American children born in the United States (to Haitian parents) that are influenced not only by Kreyol, but by English as well. This study will specifically look at the production of vowels. Vowels were chosen, in part, because their relative “steady-state” quality holds valuable information that listeners use to help interpret the speech signal (Pickett, 1999). Another reason vowels were chosen is because they are early developing sounds that aren’t typically misarticulated in children (McLeod & Bleile, 2004; Bernthal & Bankson, 1993).

Vowels are described acoustically by their location in the vowel space, which is guided by the position/shape of the articulators during production. The first formant (F1) and the second formant (F2) provide information regarding vowel quality. F1 corresponds to tongue height (Ladefoged, 1996). As the tongue rises, F1 decreases; as the tongue lowers, F1increases. F2 corresponds to tongue advancement. As the tongue moves forward, F2 increases; as the tongue moves backward, F2 decreases (Mosser, 1999; Pickett, 1999; Rosner & Pickering, 1994; Klatt, 1976).

This study will use the acoustic parameters of F1 and F2 to provide an objective description of speakers‘ vowel spaces as they produce words containing American English and Haitian Kreyol vowels. The following questions will be addressed:

1. What is the acoustic representation of the Haitian Kreyol vowel space when spoken by 4- and 5-year olds of Haitian descent?

2. What is the acoustic vowel space of American English vowels spoken by 4- and 5-year olds of Haitian descent?

3. How does this English acoustic space produced by Haitian speakers compare to that of native speakers of American English that live in the same region and are of similar ages?

METHODS

Subjects

Ten speakers of Haitian descent (6 male, 4 female) and eight non-Haitian speakers (4 male, 4 female), ages 5-6 years participated in the study (Table 2). Participants were recruited from kindergarten classes at 2 private church schools in Fort Lauderdale (Broward County), Florida. The schools were located on the border of Lauderhill and Plantation, Florida, which is an area with a large concentration of Haitian Americans.

Table 2: *Participant Description*

Group	**Male**	**Female**	**Total (N)**
Haitian American Monolingual (HAM)	3	2	5
Haitian American Bilingual (HAB)	2	3	5
Non-Haitian (NH)	4	4	8

Of the Haitian American children, five were monolingual English speakers and five were bilingual. Three of the five bilingual subjects were reported to have learned Kreyol first. Four of the five were reported to be in a home where Kreyol was spoken (Table 3). All non-Haitian children in the study were monolingual English speakers.

Table 3. *Information from language use questionnaire: Language input for each Haitian American participant*

Subject	**Language Status**	**Language Spoken in the Home**	**Language Learned First**	**Exposed to Kreyol Daily**
3	Bilingual	Unknown	Unknown	Unknown
4	Bilingual	Kreyol	Kreyol	Yes
5	Monolingual	English	English	Yes
6	Bilingual	Kreyol/English	Kreyol	Yes
7	Monolingual	English	English	No
8	Monolingual	English/Kreyol	Kreyol	Yes
10	Monolingual	Unknown	Unknown	Unknown
17	Monolingual	English	English	Yes
18	Bilingual	Kreyol/English	Kreyol/English	Yes
19	Bilingual	Kreyol	Kreyol	Yes

Stimuli

Two sets of picture stimuli were presented to elicit word productions in both English and Kreyol. Table 4 provides a list of stimulus items. The English stimulus set contained 18 pictures, whereas the Kreyol stimulus set contained 20 pictures. Both sets of picture stimuli were flashcards of everyday household/play items gathered from Baby Bumblebee (http://www.babybumblebee.com).

Table 4. *Stimulus items* (www.babybumblebee.com; www.kreyol.com; Savain, 1999)

English	**IPA**	**Kreyol**	**IPA**
Boot	but	Bouch (Mouth)	buʃ
Spoon	**spũn**	Boul (Ball)	bul
Baby	bebi	Zye (Eye)	zye
Train	**trẽn**	Pye (Foot)	pye
Boat	bot	Chapo (Hat)	ʃɑpo
Comb	**kõm**	Dlo (Water)	dlo
Tree	tri	Liv (Book)	liv
Key	ki	Bis (Bus)	bis
Pig	pɪg	Flé (Flower)	flɛ
Fish	fɪʃ	Chez (Chair)	ʃɛz
Dress	drɛs	Mato(Hammer)	mɑto
Leg	lɛg	Tab (Table)	tɑb
Hand	**hæ̃nd**	Ból (Bowl)	bɔl
Cat	kɑt	Póm (Apple)	pɔm
Clock	klɑk	**Chen (Dog)**	**ʃɛ̃**
Car	kɑr	**Nen (Nose)**	**nɛ̃**
Cup	kʌp	**Elefan (Elephant)**	**ɛlɛfɑ̃**
Duck	dʌk	**Zoranj (Orange)**	**zorɑ̃ʒ**
		Lion (Lion)	**liõ**
		Avyon (Plane)	**ɑvyõ**

*Stimuli in **bold** indicate nasal vowels in Kreyol or nasalized vowels in English

Elicited words contained the following vowel sounds: / i, ɪ, e, ɛ, æ, ʌ, u, o, ɑ, ɔ, ẽ, ɑ̃, õ/. Oral vowels, in general, exist in the vowel inventory of both English and Kreyol. Although nasal vowels occur in English, they are allophones of oral vowels and do not represent separate and distinct phonemes. However, in order to maintain balance in the stimuli, English words that have features of nasal assimilation were included in the stimulus set. Stimuli were presented in three syllabic formats. In English, stimulus items were in "Consonant Vowel Consonant" (CVC) or CCVC format only since English lax vowels cannot occur in open syllables. However, in Kreyol, the labels for the target vowels can occur in CV, CVC, or CVCV syllables.

Procedures

Subjects were tested in a quiet room. Stimuli were presented to subjects using flash cards. Each child was asked to provide a label for each given stimulus item in English and/or Kreyol (see the Appendix for elicitation script). When a child was unable to provide a label, a model was provided and the child was asked to repeat the label. Each subject participated in a practice session in order to become familiar with the procedures.

Subjects wore a head-mounted Platronics DSP-400 headset with a noise-canceling microphone. Subjects' productions were recorded directly onto a computer hard drive as a Quick Time audio file (.mov file) using a sampling rate of 44.1 kHz with 16-bit quantization. Tokens were then exported as a .wav file using the same sampling/quantization rate and saved on disk.

Acoustic Analysis

Before segmentation, samples underwent a noise reduction process to reduce ambient noise in the signal. A portion of each sample (which contained ambient noise only) was taken out to create a noise profile. This profile was then applied to the subject's entire signal. This "noise reduction" process was implemented for each subject. Once each sample went through the process, it was saved to a disk and down-sampled to 11.025 kHz for spectrographic analysis.

Vowel onsets and offsets were located manually (using the waveform and spectrogram as a reference). Determination of vowel onset location was made as follows:

1. For vowels preceded by stop consonants: Vowel onsets were marked just after the release of the consonant (and at the beginning of voicing for a particular vowel). Vowel offsets were marked at the beginning of the closure for the final consonant following the target vowel.
2. For vowels preceded by the liquid /l/: The vowel onset was marked following elimination of the spectral zero (produced by alveolar contact) and increased energy in the F2-F3 frequency range.
3. For vowels preceded by the liquid /r/: Vowel onsets were marked at the point where F3 was raised and separated from F2 (i.e., the point at which the sound was no longer rhotacized).
4. For vowels followed by the liquid /r/: Vowel offsets were marked at the point where F3 was lowered to a frequency close to F2 (i.e., onset of rhotacticization).
5. For vowels followed by nasals /m, n/ with the exception of Kreyol nasal vowels: Vowel offsets were marked at the drop in amplitude energy at the third formant.

The onsets and offsets were used to calculate the overall vowel duration. Formant frequency values were then extracted manually with 98% pre-emphasis using LPC with a 450 Hz

bandwidth and a Hamming window. F1 and F2 frequencies were measured at three temporal points within the duration of the vowel:

1. 20 ms from vowel onset
2. at vowel midpoint
3. 20 ms from vowel offset

A two-way ANOVA (with group and gender as the between-subject factors) was performed on midpoint F1 and F2 values to determine if significant differences existed between Haitian American monolingual (English) and bilingual (Kreyol/English) speakers during the production of Kreyol and English vowels. Only group differences will be discussed here.

RESULTS

In order to describe the vowel production skills of bilingual Haitian American children, the following questions were addressed:

1. *What is the acoustic representation of the Haitian Kreyol vowel space when spoken by 4 and 5 year olds of Haitian descent?*

Mean F1 and F2 values (and standard deviations) are provided in Table 5. When plotted, formant frequency values indicated that the basic Kreyol vowel space is triangular with three ”point” vowels, /i, u, ɑ/ (Figure 1). When these values were compared across the two Haitian American groups (Haitian American monolingual and Haitian American bilingual speakers), differences were not statistically significant.

Table 5. *Mean F1 and F2 measures (in Hz) for Kreyol vowels produced by all Haitian subjects*

Vowel Type	F1	F2
õ	547 (186.0)	1519 (271.0)
ɛ̃	783 (121.6)	2513 (234.1)
ɑ̃	955 (172.3)	1860 (201.8)
i	406 (66.3)	2928 (270.1)
u	481 (49.6)	1191 (113.9)
e	662 (32.3)	2508 (157.3)
ɑ	1091 (154.1)	1760 (125.8)
o	623 (84.9)	1241 (113.0)
ɛ	731 (65.9)	2403 (83.9)

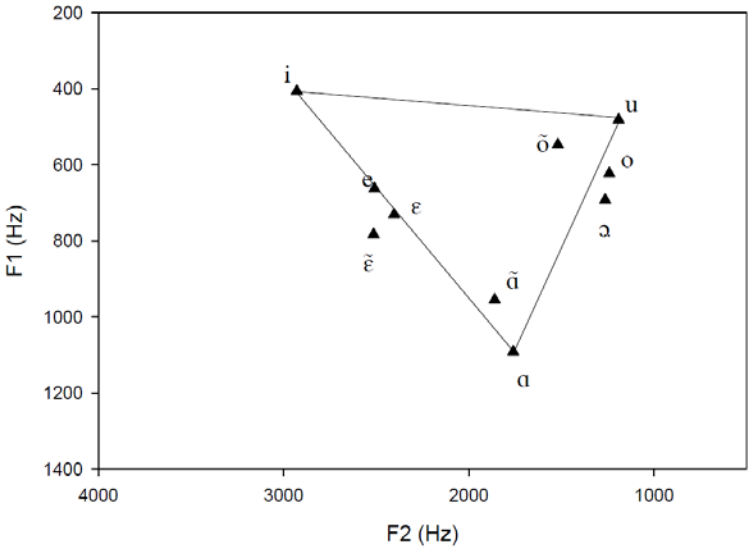


Figure 1. Kreyol vowel space for Haitian American subjects

ɔ	692 (141.1)	1263 (206.3)

2. *What is the acoustic vowel space of American English vowels spoken by 4-and 5-year olds of Haitian descent?*

Mean group F1 and F2 values for English vowels produced by Haitian American subject are provided in Table 6. The corresponding vowel space (in Hz) for Haitian American subjects can be found in Figure 2. The overall shape of the vowel space is consistent with previous studies that describe the English vowel space (Peterson and Barney, 1952; Hillenbrand, Getty, Clark & Wheeler, 1995; Lee, Potiamianos, & Narayanan, 1999). The vowel space of Haitian American subjects contains four "point" vowels (i, u, ɑ, ae), three mid front vowels (ɛ, ɪ, e), one middle vowel (ʌ), and one mid back vowel, (o).

Table 6. *English F1 and F2 values (in Hz) for all Haitian American subjects*

Vowel	F1	F2
æ	1036 (99.2)	2295 (148.7)
e	543 (83.4)	2832 (170.4)
ɛ	790 (69.6)	2330 (132.4)
i	406 (42.8)	3297 (188.4)
ɪ	592 (59.6)	2677 (163.0)
o	670 (76.0)	1237 (194.8)
u	441 (64.3)	1117 (245.5)
ɑ	1037 (83.3)	1529 (112.5)
ʌ	855 (50.9)	1694 (236.6)

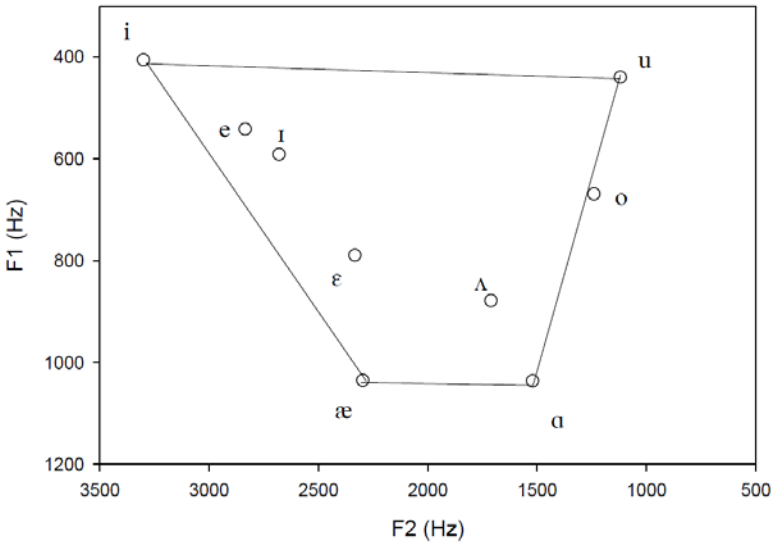


Figure 2. English vowel space (in Hz) for all *Haitian American subjects*

When F1 and F2 values for HAM and HAB were compared, results indicated significant main effects of group on F1 values for /o/ [$F(1, 6)=9.9$, $p<.05$, $\square^2=.623$]. HAM speakers produced /o/ lower in the vowel space (closer toward /ʌ/), whereas HAB speakers produced /o/ higher in the space, closer toward /u/ (Figure 3). This difference in production could be attributed to possible diphthongization on the part of bilingual speakers. Or it could have been an effect of consonant co-articulation and differences in production of the final consonants. It should be

noted that HAM and HAB speakers produced all other English vowels with similar F1 and F2 values (Table 7).

Table 7. *Mean English F1 and F2 values (in Hz) broken down by language status: Haitian American monolingual (HAM) and Haitian American bilingual (HAB) subjects*

Vowel	HAM		HAB	
	F1	F2	F1	F2
æ	1018 (117.0)	2320 (125.0)	1055 (67.9)	2269 (180.2)
e	497 (83.8)	2807 (158.1)	588 (58.9)	2857 (196.8)
ɛ	780 (81.0)	2317 (173.9)	800 (63.9)	2343 (93.8)
i	422 (43.1)	3217 (227.0)	391 (40.9)	3378 (110.4)
ɪ	590 (86.1)	2679 (207.7)	594 (23.9)	2675 (129.0)
o	726 (45.6)	1352 (201.1)	614 (55.0)	1122 (108.1)
u	451 (77.0)	1190 (286.5)	431 (55.9)	1043 (200.1)
ɑ	1023 (92.2)	1519 (110.0)	1052 (81.3)	1539 (127.0)
ʌ	833 (62.2)	1571 (155.8)	877 (27.8)	1818 (252.2)

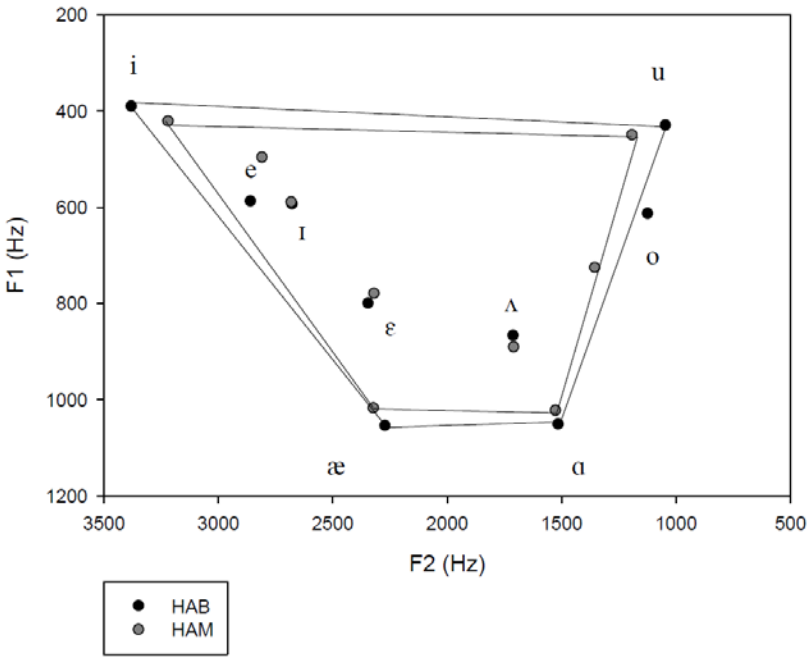


Figure 3: English vowel space for Haitian American monolingual and bilingual speakers.

3. *How does this English acoustic space produced by Haitian speakers compare to that of native speakers of American English that live in the same region and are of similar ages?*

In order to determine if Haitian American speakers produced English vowels differently than a Non-Haitian speaker, their productions were compared with Non- Haitian (NH) peers from the same region. Formant values (in Hz) for English vowels as a function of group can be found in Table 8.

Table 8. *Mean English F1 and F2 Values (in Hz) for Haitian American and Non-Haitian Speakers.*

Vowel	Haitian		Non-Haitian	
	F1	F2	F1	F2
æ	1036 (92.2)	2295 (148.7)	1114 (176.8)	2382 (191.9)
e	543 (83.4)	2832 (170.4)	517 (71.5)	2859 (243.4)
ɛ	790 (69.6)	2330 (132.4)	772 (67.8)	2443 (100.2)
i	406 (42.8)	3297 (188.4)	428 (46.9)	3299 (211.7)
ɪ	592 (59.6)	2677 (163.0)	567 (34.8)	2759 (189.6)
o	670 (76.0)	1237 (194.8)	635 (90.0)	1201 (181.8)
u	441 (64.3)	1117 (245.5)	456 (111.3)	1281 (165.6)
ɑ	1037 (83.3)	1529 (112.5)	1025 (89.3)	1510 (171.2)
ʌ	855 (50.9)	1695 (236.6)	903 (94.8)	1683 (131.1)

At first glance, the F1 values for Haitian American speakers‘ /æ/ and /u/were observed to be higher than their Non-Haitian counterparts. However, ANOVA results indicated that differences in F1 between the groups were not significant. This could be attributed to high variability during production. ANOVA results also indicated no significant differences in F2 values between Haitian American and non-Haitian speakers. Overall, both Haitian American and non-Haitian speakers produced English vowels with similar tongue height and “forwardness” (Figure 4).

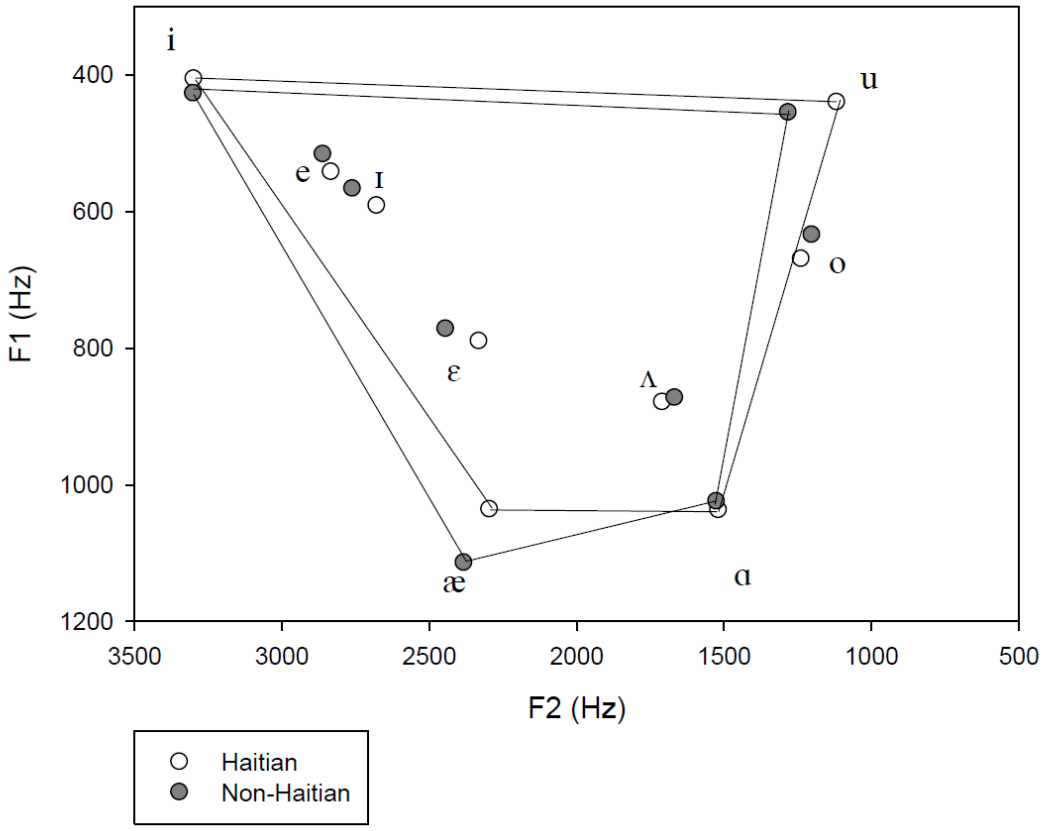


Figure 4. English vowel space for Haitian American and non-Haitian speakers

Because of their status as monolingual English speakers, HAM speakers were factored out and Haitian American bilingual speakers' productions were compared to their non-Haitian (NH) peers from the same region. Formant values (in Hz) for English vowels as a function of group can be found in Table 9. ANOVA results did not indicate significant F1 and F2 differences between groups. Overall, both Haitian American bilingual and Non-Haitian speakers produced English vowels with similar tongue height and "forward/backward" tongue movement (Figure 5).

Table 9. *Mean English F1 and F2 values (in Hz) for Haitian American bilingual and non-Haitian speakers.*

Vowel	Haitian American Bilingual		Non-Haitian	
	F1	F2	F1	F2
æ	1055 (67.9)	2269 (180.2)	1114 (176.8)	2382 (191.9)
e	588 (58.9)	2857 (196.8)	517 (71.5)	2859 (243.4)
ɛ	800 (63.9)	2343 (93.8)	772 (67.8)	2443 (100.2)
i	391 (40.9)	3378 (110.4)	428 (46.9)	3299 (211.7)
ɪ	594 (23.9)	2675 (129.0)	567 (34.8)	2759 (189.6)
o	614 (55.0)	1122 (108.1)	635 (89.9)	1201 (181.8)
u	431 (55.9)	1043 (200.1)	456 (111.3)	1281 (165.6)
ɑ	1052 (81.3)	1539 (127.0)	1025 (89.3)	1510 (171.2)
ʌ	877 (27.8)	1818 (252.2)	903 (94.8)	1683 (131.0)

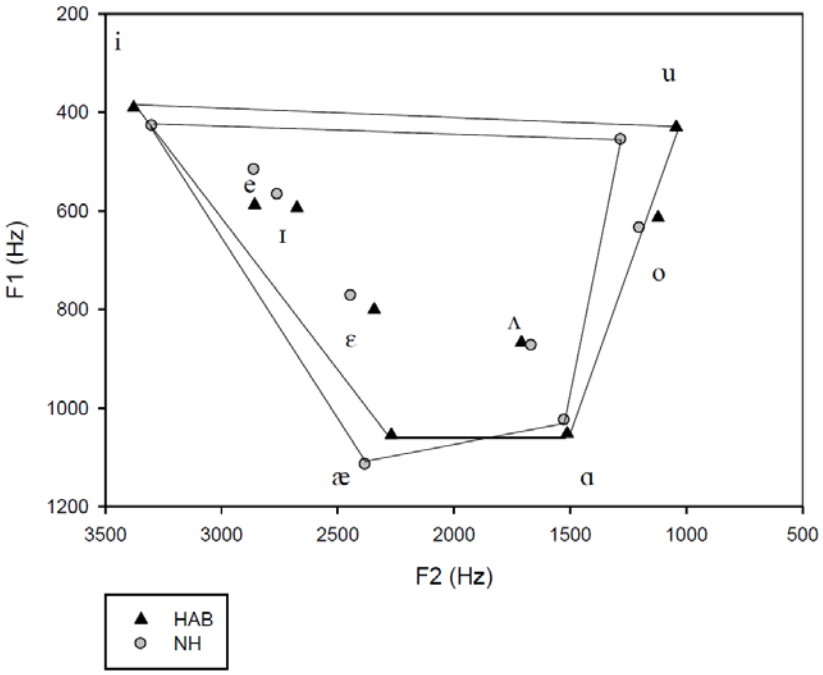


Figure 5. English vowel space for Haitian American bilingual and non-Haitian speakers

DISCUSSION

This study was designed to examine how Haitian (bilingual) speakers of Kreyol and English produce the sounds in the two languages. Specifically, this research provided an acoustic description of Kreyol and English vowels produced by monolingual (English) and bilingual (Kreyol/English) children of Haitian descent.

Results of this study revealed that the acoustic description of the Kreyol vowel space produced by this group of Haitian American speakers (e.g., residents of South Florida) is consistent with non-acoustic adult descriptions (Tinelli, 1981). With South Florida being one of the three states that report the largest Haitian American population, these results are to be expected. Similarities in phonemic repertoire could have also contributed to these results.

Another issue to consider is the fact that there were no monolingual Kreyol speakers in this study. It is not clear if these productions are true representations of Kreyol (not influenced by English). Having data on how monolingual Kreyol speakers produce Kreyol vowels would provide baseline measurements that would allow for a better monolingual/bilingual comparison.

When Haitian American speakers produced English vowels, acoustic analyses revealed that the English vowel space produced was no different than the vowel space of non-Haitian native English speakers. This was the case when both monolingual and bilingual Haitian speakers were compared to non-Haitian native English speakers from the same geographical area. This suggests that bilingual speakers as young as 5 years old can produce the vowel sounds of their second language as a native speaker. These overall results support the SLM, which indicates the influence of AOA on native- like production (Flege, 1999). The fact that the young bilingual speakers in this study appear to be able to differentiate between Kreyol and English vowels (their native and second language), leads one to wonder if these same results would exist during the production of consonant sounds. Examining consonant production differences between bilingual and monolingual speakers would be a useful follow-up to this research. If the same results are seen, it might be possible to test the articulation skills of a bilingual speaker of this particular age in the second language only.

With more recent descriptions of vowels looking at how the spectral features of the vowel change over the length of the vowel (Fox & McGory, 2007; Jacewicz, Fox, & Salmons, 2006; Fox, Jacewicz, & Salmons, 2006), future research should investigate dynamic spectral change to see if similarities in formant frequencies across groups continue to be evident.

Further research should also investigate if differences in production occur as a function of age. Conducting a cross-sectional study that investigates differences in vowel characteristics of Haitian American speakers of different ages and different ages of arrival (AOA) in the United States would help determine the critical age for bilingual speakers producing vowels as a native speaker.

ABOUT THE AUTHORS

Stacey Wallen is a speech-language pathologist and Assistant Clinical Professor at Georgia State University. Her current research interests include examining differences in the perception and production of phonemes as it relates to culturally/linguistically different groups. She has a Ph.D. in speech and hearing science from The Ohio State University.

Robert A. Fox is a Professor and Chair of the Department of Speech and Hearing Science at The Ohio State University. He is best known for his contributions to speech perception, especially for his work on the perceptual structure of vowels and diphthongs. He has also studied behavioral aspects of speech perception such as changes in listener response as a function of human aging and second language learning as well as dialectal variation.

Correspondence concerning this article should be addressed to Stacey Wallen, Department of Educational Psychology and Special Education, Georgia State University, Atlanta, GA 30303. Phone: (404) 413-8343. Email: swallen@gsu.edu

REFERENCES

http://www.babybumblebee.com

Bernthal, J. & Bankson, N. (1993). *Articulation and phonological disorders.* New Jersey: Prentice Hall.

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

Flege, J. E. (1995). Second-language speech learning: Findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233-273). Timonium, MD: York Press.

Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.). *Second Language Acquisition and the Critical Period Hypothesis* (pp. 101- 131). Mahwah, NJ: Lawrence Erlbaum Associates.

Flege, J., MacKay, I, & Meador, D. (1999). Native Italian speakers‘ perception and production of English vowels. *Journal of the Acoustical Society of America, 106,* 2973-2987.

Fox, R., Jacewicz, E. & Salmons, J. (2006). Prosodically induced phonetic variations in vowels: A source of language change? In H. I. Aronson, F. L. Dyer, V. A. Friedman, D. S. Hristova, & J. M. Sadock (Eds.), *The Bill Question*, (pp.87-110). Bloomington, IN: Slavica.

Fox, R. A. & McGory, J. (2007). Second language acquisition of a regional dialect of American English by native Japanese speakers. In O.S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning*, (pp. 117-134). Amsterdam: John Benjamins Press.

Goldstein, B. (2001). Transcription of Spanish and Spanish-influenced English. *Communication Disorders Quarterly, 23,* 54-66. Retrieved July 9, 2005, from Questia database, http://www.questia.com.

Gorman, B.K., & Kester, E.S. (2003). Spanish-influenced English: Typical phonological patterns in the English language learner. Online continuing education course. http://www.bilingusitics.com.

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

Jacewicz, E., Fox, R. A. & Salmons, J. (2006). Prosodic prominence effects on vowels in chain shifts. *Language Variation and Change 18,* 285-316.

Kindler, A. (2002). *Survey of the States' Limited English Proficient Students and Available Educational Programs and Services.* 2000-2001 Summary Report.

http://www.ncela.gwu.edu/files/rcd/BE021853/Survey_of_the_States.pdf

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America 59*, 1208-122l

Ladefoged, P. (1996). *Elements of Acoustic Phonetics*. Chicago: University of Chicago Press.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*, 1455-1468.

McLeod, S. & Bleile, K. (2004). Typical speech development summary. A resource summarizing research on typical speech development. Retrieved from http://athene.riv.csu.edu.au/~smcleod/ASHA03McLeod&Bleile.pdf

Mosser, G. (1999). Forming formants. In Scientific American: The Throat Singers of Tuva. Retrieved from http://www.sciam.com/1999/0999levinbox3.html.

Muysken, P. & Veenstra, T. (1995). Haitian. In Arends, J., Muysken, P., and Smith, N. (Eds). *Pidgins and creoles: An introduction*. Philadelphia, PA: John Benjamins Publishers.

Peterson, G. & Barney, H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*, 175-184.

Pickett, J. (1999). *The acoustics of speech communication*. Boston: Allyn and Bacon.

Portes, A. & Rumbaut, R. (2001). *Legacies: The story of the immigrant second generation*. New York: Russell Sage Foundation.

Regan, J. (2003). Haiti's first language still running second. Retrieved from http://www.hartfordhwp.com/archives/43a/538.html

Rosner, B. & Pickering, J. (1994). *Vowel perception and production*. New York: Oxford University Press.

Savain, R. (1999) *Haitian Creole in ten steps*. Rochester, VT: Schenkman Books.

St. Fort, H. (2000). *What is Haitian Creole?* Retrieved from http://www.ahadonline.org/eLibrary/creoleconnection/Number20/haitiancreole.htm

Tinelli, H. (1981). *Creole phonology*. New York: Mouton.

APPENDIX

Elicitation Script

I'm going to show you some pictures. Your job is to tell me what each picture is. Each time I show a picture I'm going to ask "What is this?" I want you to answer "It's a _____" and say the name of the picture. So if I show you a picture of a bird (hold up the picture of the bird) **and say "What is this?" I want you to say "It's a bird."**

Let's try some for practice.

(Hold up example #1—a picture of a shirt) What is this? *(Wait for response. If child answers "It's a shirt" then reply* ***"You're right it's a shirt, good job"*** and continue to stimulus items).

(If child answers "shirt", praise child for correct answer then model the desired response ***"You're right, it's a shirt. Can you say "It's a shirt.?"*** Wait for child to repeat the desired response then move to example #2—a picture of a cookie—and repeat process).

(If child appears unable to label the picture, provide a verbal cue (i.e. ***"It's a sh__".***) and wait for child to produce label. If the child is unable to name the picture after verbal cue, provide the label and ask child to repeat).

Pearson, P., Pickering, L., & Da Silva, R. (2011). The impact of computer assisted pronunciation training on the improvement of Vietnamese learner production of English syllable margins. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 169-180), Ames, IA: Iowa State University.

THE IMPACT OF COMPUTER ASSISTED PRONUNCIATION TRAINING ON THE IMPROVEMENT OF VIETNAMESE LEARNER PRODUCTION OF ENGLISH SYLLABLE MARGINS

Pamela Pearson, Georgia State University

Lucy Pickering, Texas A & M, College Station

Rachel Da Silva, Georgia State University

In this study we investigate the utilization of computer assisted pronunciation training to the issue of syllable margin production by Vietnamese learners of English. Thirteen intermediate Vietnamese students were recruited from an intensive English program. A pre-test was administered to elicit student performance on the production of syllable margins (e.g. onsets and codas). Treatment materials comprised a second set of items featuring audio files recorded by a native speaker of American English and used as the target models for training. Eight 30-minute tutoring sessions were conducted in which students used a KayPentax Computerized Speech Laboratory to compare their self-produced spectrograms with the prerecorded target spectrograms. Post-test measures indicated some significant movement toward more target-like production of syllable margins. This paper describes the pedagogical methods and materials used in this study with an emphasis on word list creation, design of intervention and coding and analysis. We outline the benefits of this training for pronunciation improvement including the positive reactions of both teachers and students to the use of this teaching technique.

Among practitioners of ESL/EFL, certain pronunciation challenges are recognized as in some way intractable. Not that they cannot be improved, but that improvement may be rare and comes at a high cost in terms of time and effort for both learners and teachers. One such area is the production of syllable margins in the speech of Vietnamese learners (Hansen, 2006). One relatively recent approach to addressing obstinate pronunciation issues is the use of multimodal or visual feedback through the use of speech visualization technology (Hardison, 2004, 2005; Levis & Pickering, 2004). Computer assisted pronunciation training (CAPT) (Levis, 2007) encompasses a range of tools and techniques which include the use of spectrograms, pitch contours and statistical information on acoustic features of oral production. What they have in common is an *analytic-linguistic* approach to pronunciation training as opposed to the more traditional *imitative-intuitive* approach (Celce-Murcia et al, 1996). Initial investigation of these techniques suggests that they may enhance auditory-only instruction thus we conducted a pilot study to examine the effect of CAPT on the syllable margin production in Vietnamese language learners.

LITERATURE REVIEW

Syllable structure transfer in L2 acquisition has been frequently documented in the SLA literature (see Leather & James, 1991 for a review). Although the investigation of the phonological development of Vietnamese learners of English (VLEs) is rare, early studies (Benson, 1988; Sato, 1984, 1985) emphasize the role of L1 transfer in syllable margin production leading to consonant cluster reduction and the differential production of word-initial and word-final clusters. This is reiterated by Osburne (1996) in her case study of a Vietnamese learner. In two interviews over a span of six years, she demonstrates that cluster reduction is not random but rather regular and "influenced subtly by the expectations of L1 syllable structure" (Osburne, 1996, p. 175). Vietnamese lacks consonant clusters (Avery & Ehrlich, 1992; Honey, 1987) and has a tendency toward open syllables, which may contribute to word ending deletion (Hwa-Froelich et al., 2002; Osburne, 1996).

In addition to the investigation of linguistic constraints including L1 transfer, Hansen (2006) examines socio-affective factors including social identity, gender, context of use, and investment on the acquisition of syllable onsets and codas of a Vietnamese married couple who begin as newcomers to the US. During her 10-month study, Hansen finds little overall change in terms of ultimate acquisition (approximately 85% for onsets and 40% for codas), however she documents some emerging production modifications toward targetlike production that may lead to change over a longer period of time. These studies present a somewhat gloomy picture for VLEs as traditional methods of instruction do not appear to have a significant impact on development. Syllable structure may become a "stabilized" aspect of Vietnamese interlanguage phonology (Han, 2004; Long, 2003).

CAPT techniques in pronunciation instruction represent a novel approach only recently widely accessible to teachers and students through downloadable freeware such as *Praat* (Boersma & Weenink, 2005). Speech visualization feedback has been utilized to examine segmental and suprasegmental issues in pronunciation (Hardison, 2004; Hirata, 2004; Lambacher, 1999; Levis & Pickering, 2004). Lambacher (1999) uses "electronic visual feedback" with Japanese learners to practice the production of English consonant contrasts: oral and nasal stops, liquids, and fricatives. Although he claims that the students made "articulatory adjustments" based on direct comparison with the L1 models (p. 145), he provides no empirical evidence as to the actual acquisition of these contrasts. Hardison (2005) used visual displays of learners' pitch contours to provide contextualized prosody training and found significant improvement following training in addition to positive reactions from learners regarding the use of visual feedback. To our knowledge, no one has used this kind of feedback to explore the acquisition of English syllable margins.

In light of the intelligibility problems displayed by our VLEs, many of whom reported having studied English for five years or more, we undertook the following study to investigate if we could destabilize VLE syllable structure through an approach designed to cultivate apperceived input (Gass & Selinker, 2001). In this form of CAPT, spectrograms of the learners' spoken production were used to provide visual feedback.

METHOD

Participants

In total, 13 students participated in the study: five (4 = F; 1 = M) during the summer of 2008, and eight (2 = F; 6 = M) in the spring of 2009. The mean age of the participants was 23 (range: 19 to 28 years); the mean amount of time they had spent studying English in the US was five months (range: 0.5 months to one year); and the mean length of time they had spent studying English in Vietnam was five years (range: two years to eight years). At the time of the intervention, each student was enrolled in an Intensive English Program (IEP) at a major university in the Southeastern United States. The students were taking 15 semester hours of classes at the IEP including a three-hour course in Oral Communication. Participant proficiency level was classified as pre-intermediate (n = 5), intermediate (n = 6), or upper intermediate (n = 2) based on the results of an in-house placement exam, which included a proficiency interview.

Materials

Pre- and post-tests were constructed to elicit the participants' performance with regard to the perception and production of syllable margins (i.e., onsets and codas.) Each test comprised (a) 30 multiple-choice items to test perception, (b) 55 items to test production, and (c) a semi-structured speaking task[1]. To avoid any potential interference caused by suprasegmental or literacy issues, all items were monosyllabic words. Decisions regarding which syllable margins to incorporate were made according to their appearance in the lower tiers of accuracy percentiles on similar tasks with VLEs, as reported by Hansen (2006, pp. 66, 78-79).

Our final word list comprised 10 singleton onsets, 10 two-member onsets, and five three-member onsets (the maximum possible in English); it also contained 10 items in each of the following categories: single, double and triple codas (see Appendix A for complete listing). To ensure internal consistency and equivalence of forms, items on the pretest were repeated but randomized on the post-test.

A second set of items featuring the same onsets and codas was created for use in the actual treatment in order to minimize the practice effect (Appendix A). From this second set, the materials required to administer the treatment were developed. WAV audio files for each item were digitally recorded by one of the researchers, a native speaker of American English, and filed for use as spectrogram models, and a packet of student-use materials depicting items in both print and image form was created (Appendix B).

Procedures

The treatment was carried out in either instructor-pair or instructor-small group (n = 3) tutoring sessions using a Kay Pentax Computerized Speech Laboratory (CSL) Model 4500 equipped with Real-Time Spectrogram software. Eight 30-minute sessions (two per week for four weeks) were conducted. The first and last meetings were reserved for testing, and there were six tutoring sessions in between. A breakdown of the session activities schedule follows in Table 1.

[1] These data are not examined here

Table 1. *Session Activity Summary*

Week	Session	Activity	Session	Activity
1	1	Intro to the study & CSL; IRB consent; pretest	2	Single onsets
2	3	Double onsets	4	Triple onsets
3	5	Single codas	6	Double codas
4	7	Triple codas	8	Post-test

During each session, students were provided with speech visualization feedback using the CSL (see Figure 1 below). Students were shown a model of the target word and trained to recognize the components of the spectrogram. They then produced the word and analyzed their self-produced spectrograms in light of the model spectrograms. Sessions were driven by an emphasis on mastery, meaning that students were allowed to record items as many times as necessary in order to produce spectrograms with which they were satisfied.

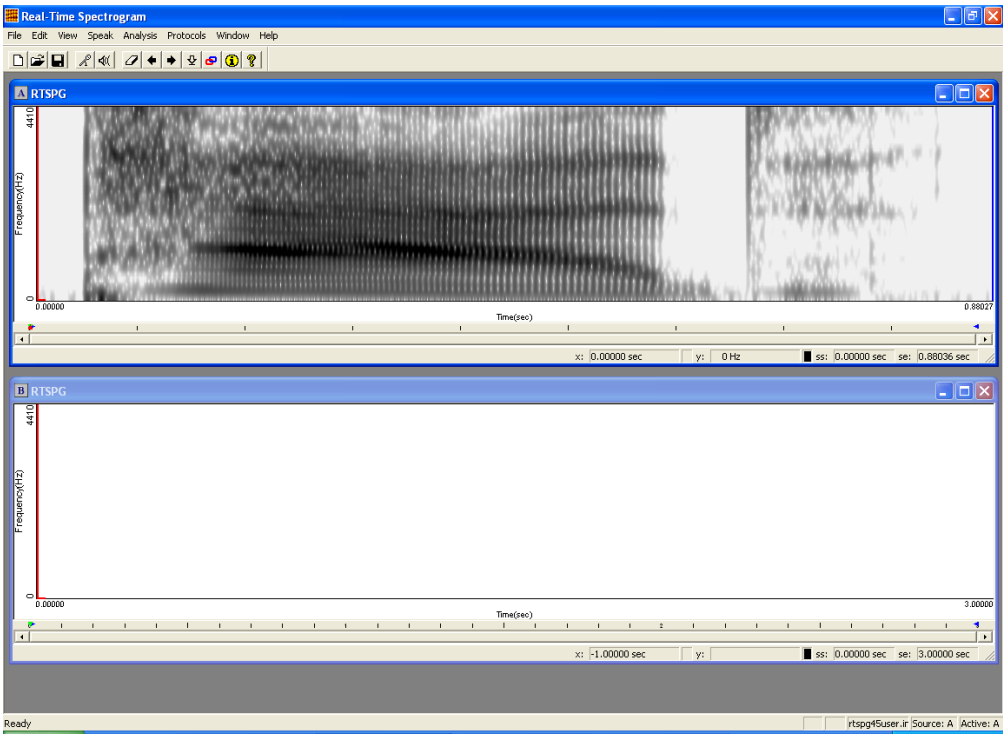


Figure 1. Sample screenshot of *pad* in Real-Time Spectrogram software

Data Analysis

In order to determine whether positive gains occurred from Time 1 to Time 2, paired samples *t*-tests were run in *SPSS 15.0* (SPSS Corporation, 2006). Perception data was entered as either 'correct' or 'incorrect'. However, due to a high group mean on the pretest, and consequent lack of a normal distribution of the scores, a *t*-test was not a viable analysis to perform for the measure of perception and simple percentages were used instead. Production data were coded prior to analysis. Coding forms were created for both versions of the test (see Appendix C) which allowed for both the dichotomous coding of items – targetlike or non-targetlike – and also for the diagnosis of non-targetlike productions according to categories used by Hansen (2006): deletion, substitution, epenthesis and other. In order to receive a coding of 'targetlike', an item had to be judged as such by both raters. Using Cohen's Kappa, interrater reliability across the entire data set was .71. In simple percentages, inter-rater reliability was 88.7% (agreement = 634/715 items) for the pretest, and 82.5% (agreement = 590/715) for the post-test data.

RESULTS

With regard to the perception tests, the participants collectively scored 338 out of a total of 390 multiple-choice items resulting in an 86.7% rate of accuracy on the pretest. As a group, they performed marginally higher on the post-test, making 36 errors on the 390 items resulting in a 90.8% rate of accuracy with an overall rate of improvement of 4.1%. In contrast to the high rate of accuracy on the perception pretest, on the production pretest, the participants produced only 255 targetlike syllable margins (out of a possible 715) for a combined accuracy rate of 35.7%. On the post-test measure, however, students demonstrated a considerably higher rate of accuracy. Overall, the rate of accuracy rose to 52.2% (373 targetlike productions on 715 items). A paired samples *t*-test conducted with this data found a statistically significant increase of group means (from 19.6 on the pretest to 28.7) ($p = .01$ where $p < .05$).

Items were also clustered by syllable margin categories and ranked by percentage of targetlike production: single onsets (80%), double onsets (62%), triple onsets (44%), single codas (42%), double codas (40%), and triple codas (28%). The data show that codas proved to be more troublesome for the VLEs than onsets in both perception and production tests. These results are highly consistent with the findings of Hansen (2006).

DISCUSSION

This study suggests that pedagogical intervention may be valuable in terms of improving VLEs' production of syllable margins at least in the short term. Following tutoring sessions employing CAPT methods, early post-tests show a significant difference in overall accuracy of production of forms produced in a word-list format. In post-test production, raters reported more difficulty in assessing targetlike vs. non-targetlike forms resulting in a slightly lower agreement rate. We suggest that these changes in production show that syllable structure can be "destabilized" in the IL of these learners. Clearly, it will be important to expand the treatment period, vary tasks, and allow for continued delayed post-testing in our follow up work in order to determine the applicability and duration of the effect.

In addition, in further tests we will reconsider some of our word list choices. Four of the items were produced in a non-targetlike manner by all five of the participants at both Times 1 and 2: *fudge*, *teethe*, *cloths*, and *minced*. We realized that word recognition (e.g., *cloths* might look like the more frequent *clothes*) could play an important role in whether the students were able to achieve targetlike production (Venkatagiri & Levis, 2007). In the future, we will control for frequency by consulting a word frequency index such as the Compleat Lexical Tutor (CLT) (Cobb, 2008), which compares items to the 2000 most common word families in the Brown Corpus of American English. Of the 55 items on our word list, only 15 appear on the CLT list: *bagged*, *bands*, *bath*, *cloths*, *hid*, *hits*, *jump*, *locked*, *log*, *pray*, *pumped*, *swim*, *tap*, and *thanks*.

Finally, with regard to pedagogical implications, CAPT provided both teachers and students with an alternative form of feedback to address this particular intractable error. Through the visual aspect of the spectrogram-enhanced feedback, students were able to notice a gap in their production which was something they had not been able to do through the imitative-intuitive approach alone. Toward the end of the study, one participant commented, "I didn't know I don't say the endings of words. So, now I know, and I'm trying it when I speak in English." The teachers also responded positively to the technique. They perceived the results of the CAPT approach to be successful, and they believed they had found a more effective way to provide feedback to this particular group of learners. One teacher noted that "what [she] liked most about the CSL is that it takes the personal element out of the feedback. Instead of telling students 'no' and making them repeat over and over again, we were instead able to give them a positive goal to work towards." Although we can only report these additional findings anecdotally at this stage, we plan to include a more formal investigation of motivational benefits in our further work in this area.

APPENDIX A: Word Lists

Test items* (onsets)	Treatment items (onsets)	Test items (codas)	Treatment items (codas)
1. pat /p/	pad	26. tap /p/	lap
2. ban /b/	bat	27. hid /d/	kid
3. ten /t/	tech	28. lash /ʃ/	bash
4. zoo /z/	zoom	29. leave /v/	weave
5. shot /ʃ/	shop	30. pull /l/	full
6. vain /v/	vase	31. fudge /dʒ/	budge
7. jug /dʒ/	just	32. bath /θ/	math
8. chip /tʃ/	chick	33. teethe /ð/	breathe
9. there /ð/	they	34. locked /kt/	socked
10. think /θ/	thing	35. bagged /gd/	tagged
11. blue /bl/	bloom	36. wished /ʃt/	fished
12. play /pl/	plane	37. wind /nd/	sinned
13. crew /kɹ/	cruise	38. puffs /fs/	huffs
14. pray /pɹ/	pry	39. hits /ts/	mitts
15. swim /sw/	swish	40. cloths /θs/	moths
16. class /kl/	clam	41. pierce /ɹs/	fierce
17. three /θɹ/	throw	42. jump /mp/	bump
18. small /sm/	smog	43. clasp /sp/	asp
19. frog /fɹ/	free	44. pumped /mpt/	jumped
20. tweed /tw/	tweak	45. minced /nst/	winced
21. splash /spl/	split	46. lunched /tʃt/	bunched
22. street /stɹ/	string	47. faxed /kst/	taxed
23. spray /spɹ/	spread	48. helped /lpt/	yelped
24. scream /skɹ/	screen	49. turns /ɹnz/	burns
25. squid /skw/	square	50. flasks /sks/	masks
26. log /g/	dog	51. months /nθs/	ninths
27. robe /b/	lobe	52. bands /ndz/	hands
		53. thanks /Nks/	banks

*Accompanied by IPA transcriptions of phonological targets

APPENDIX B: Sample of Teaching Materials

APPENDIX C: Pretest production coding form sample

Name: _____

Rater: _____

Pretest Coding: Speaking

Onsets

1.**p**at	TL	non-TL: deletion, substitution, epenthesis, other
2.**b**an	TL	non-TL: deletion, substitution, epenthesis, other
3.**t**en	TL	non-TL: deletion, substitution, epenthesis, other
4.**z**oo	TL	non-TL: deletion, substitution, epenthesis, other
5.**sh**ot	TL	non-TL: deletion, substitution, epenthesis, other
6.**v**ain	TL	non-TL: deletion, substitution, epenthesis, other
7.**j**ug	TL	non-TL: deletion, substitution, epenthesis, other
8.**ch**ip	TL	non-TL: deletion, substitution, epenthesis, other
9.**th**ere	TL	non-TL: deletion, substitution, epenthesis, other
10. **th**ink	TL	non-TL: deletion, substitution, epenthesis, other

REFERENCES

Avery , P. W. & Ehrlich, S. (1992). *Teaching American pronunciation*. Oxford, UK: Oxford University Press.

Benson, B. (1988). Universal preference for open syllables as an independent process in interlanguage. phonology. *Language Learning* 38:221-235.

Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3.01) [Computer program]. Retrieved from http://www.praat.org/

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

Cobb, T. (2008). *The Compleat Lexical Tutor* [online computer software]. Retrieved from http://132.208.224.131/

Gass, S. M. & Selinker, L. (2001). *Second language acquisition: An introductory course (2nd edition)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Han, Z. (2004). Fossilization: Five central issues. *International Journal of Applied Linguistics*,14(2), 212-242.

Hansen, J. G. (2006). *Acquiring a non-native phonology: Linguistic constraints and social barriers.* London, UK: Continuum.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning and Technology, 8*(1), pp. 34-52.

Hardison, D. M. (2005). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal* 22: 175-190.

Hirata, Y. (2004). Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning*, *17*(3-4), 357-376.

Honey, P. J. (1987). Vietnamese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems (1st edition)*. Cambridge, UK: Cambridge University Press.

Hwa-Froelich, D., Hodson, B. W., & Edwards, H. T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech – Language Pathology, 11*(3), 264-273.

Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals.* Oxford, UK: Oxford University Press.

Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning*, *12*(2), 137-56.

Leather, J. & James, A. (1991). The acquisition of second language speech. *Studies in Second Language Acquisition* 13: 305-41.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27,* 1-19.

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

Long, M. H. (2003). Stabilization and fossilization in interlanguage development. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 487-535). Oxford, UK: Blackwell Publishing.

Osburne, A. G. (1996). Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics, 17*(2), 164-181.

Sato, C. (1984). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning* 34: 43-57.

Sato, C. (1985). Task variation in interlanguage phonology. In S. Gass & C. Madden (eds). *Input in second language acquisition.* Cambridge, MA: Newbury House.

SPSS Corporation (2006). *SPSS 15.0* [computer software]. Chicago, IL: SPSS Inc.

Venkatagiri, H. S. & Levis, J. M. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness*, 16(4), 263-277.

STUDENTS' AWARENESS OF SPANISH SPIRANTIZATION ALLOPHONIC RULE

Manuela González-Bueno, University of Kansas

Marcela Quintana-Lara, Universidad Arturo Prat

El propósito de este estudio es determinar si aprendices de español como lengua extranjera con diferentes niveles de habilidad lingüística tienen conciencia de la regla de espirantización de las oclusivas sonoras del español. Se espera que los resultados ayuden a determinar si esta regla se puede adquirir de forma natural, sin intervención pedagógica alguna, o si, por el contrario, se necesita entrenamiento enfocado en la forma para hacer que los estudiantes sean conscientes de la regla y consecuentemente la aprendan.

The purpose of this study is to determine the extent to which L2 Spanish learners at different levels of proficiency are aware of the spirantization rule of Spanish voiced stops. The results of this study help determine whether general exposure to the language is enough to acquire the target rule or more form-focused phonological training is needed to bring their attention to the rule and consequently learn it.

INTRODUCTION

One source of accentedness in learners of Spanish who are native speakers of English is the use of voiced stops [b, d, g] instead of the fricative [ß, ð, γ] in certain obligatory contexts for these fricative allophones (Dalbor, 1997). Spanish's spirantization allophonic rule dictates that [b, d, g] are produced as [ß, ð, γ] in rapid speech mode (andante, allegretto, and presto styles) in intervocalic position (e.g., "hada" > ['aða]) and between a vowel or a liquid (e.g., "alba" > ['alßa], "arder" > [arðer']), except in the case of [l] + [d] (e.g., "falda" > ['falda]). English does not have this rule. Although the sound [ð] is part of the English phonetic inventory as a phoneme, it has a different distribution. The sounds [ß] and [γ] rarely occur in English and are often erroneously identified as [v] and [w] respectively. Therefore, the Spanish word "haba" (['aßa]) will be pronounced as *['aba]; "hada" (['aða]), as *['ada]; and "lago" (['laγo]), as *['lago] (Table 1).

Table 1. *Lack of spirantization rule in English-accented Spanish*

Spanish word	Correct Spanish pronunciation	English-accented pronunciation
"haba"	['aßa]	*['aba]
"hada"	['aða]	*['ada]
"lago"	['laγo]	*['lago]

The purpose of this study is to determine the extent to which L2 Spanish learners at different levels of proficiency are aware of the spirantization rule of Spanish voiced stops. To that end, English-speaking students learning Spanish as an L2 were given a perception test consisting of Spanish words containing one of the obligatory Phonetics contexts for the spirantization of the target sounds [ß, ð, γ]. The spirantization rule was applied in some instances (for example, the word "dedo" was produced as [deðo]) but not in others (for example, the word "dedo" was produced as *[dedo]). Participants were asked to identify either token as "correct" or "incorrect" in the assumption that, if they were aware of the spirantization rule, they would select "incorrect" whenever they heard a stop in an obligatory context for a fricative, and vice versa.

In addition, and given that experimental studies have consistently shown a close link between production and perception in L2 learners (Llisterri, 1995), participants also read aloud a passage containing many instances of the target sounds, and their production was acoustically analyzed in order to determine if there is relationship between learners' perception and production.

Lastly, we also explore the effect of being enrolled in a Phonetics course on the perception and production of participants at the advanced level. Descriptive analysis is used to show a relationship between the test results and levels of language proficiency, and also whether a Phonetics course is taken or not. Implications for the need of form-focused phonological training are addressed.

LITERATURE REVIEW

As far back as 1941, and framed within the Markedness Theory, fricative sounds were said to be more marked than stop sounds (Jakobson, 1941). Later, when Eckman (1977) developed the Markedness Differential Hypothesis, fricatives were deemed more difficult to acquire due to their higher degree of markedness. Although the stop phones are considered the base phonemes and the fricatives their allophonic variants, in spite of the latter being more marked, fricatives are in fact much more common in Spanish than their stop counterparts (Hualde, 2005; Schwegler & Kempff, 2007).

These two facts combined represent a challenge to learners of Spanish as a foreign language. A number of studies have addressed the acquisition of Spanish spirantized allophones of voiced stops. Zampini (1994) was one of the first to directly investigate how native English-speaking L2 learners of Spanish acquire Spanish spirantization. Second and fourth-semester students participated in two tasks, one to elicit spontaneous speech and one consisting of reading a passage aloud. These tasks were designed to investigate the acquisition of fricatives and also explore the effect of speech style on their production. Zampini's results showed that all subjects produced fricatives in less than 32% of the expected instances. She also notes that the implementation of the spirantization rule might be hampered by the inability of learners to speak fast enough, but that, nevertheless, they might be aware of the rule.

In González-Bueno's (1995) study, five native speakers of English learning Spanish as an L2 were given oral proficiency tests in the form of OPI, and their productions were analyzed acoustically. In particular, instances of the obligatory contexts for [ß, ð,γ] were counted and used to determine if the spirantization rules had been applied by the students. They were found to produce fricatives about half of the time, at a higher rate than the subjects in the Zampini study.

Both Zampini (1994) and González-Bueno (1995) attribute the difficulties of L2 learners in acquiring the fricative allophones of voiced stops [ß, ð, γ] to phonemic and allophonic differences in English. Later studies also point to the difficulties of L2 learners of Spanish acquiring the spirantization process (Díaz-Campos, 2004; Elliot, 1997).

More recently, Lord (2010) analyzed the oral recordings from two groups of students participating in a study abroad program. One group had previously taken a Spanish Phonetics Course, and the other one had not. Participants read out loud a list of words and phrases in Spanish, each one containing the target sounds ([b, d, g, ß, ð, γ]). Lord (2010) concluded that explicit instruction seems to have a positive effect on the production of Spanish voiced stops, including their fricative allophones.

Given this situation, the present study was set up to contribute to this line of research by determining the extent to which L2 Spanish learners at different levels of proficiency are aware of the spirantization rule of Spanish voiced stops. In addition, and given that, in general, L2 learners do not receive explicit phonological training until they reach higher levels of language instruction, the results of this study might help determine whether general exposure to the language is enough to acquire the target rule or more form-

focused phonological training is needed to bring their attention to the rule and consequently learn it.

To that end, the present study seeks to answer three research questions:

1. At what levels of proficiency are L2 Spanish learners able to recognize whether the spirantization of Spanish voiced stops rule has been applied, as indicated by the results of the perception test?

2. At what levels of proficiency are L2 Spanish learners able to apply the spirantization of Spanish voiced stops rule, as indicated by the results of the production test?

3. Is there a correlation between L2 Spanish learners' perception and production results, as indicated by both tests?

METHOD

Participants

Eighteen native English speakers of American English without speech or hearing impairments participated in the study. They were all undergraduate students learning Spanish as a L2. At the time of the investigation, most participants were enrolled in a Spanish course at a Midwestern American University. They had an age average of 23. Two participants had a low proficiency level in a third language (French and German). Only a few participants indicated some language contact with native Spanish speakers.

Six participants had a low proficiency level; 6 participants had an intermediate proficiency level; and 6 had an advanced proficiency level. Only 2 participants (from the advanced proficiency level group) had taken a Spanish Phonetics/Phonology Course.

Stimuli

The stimuli for the perception test were 30 Spanish words containing one of the obligatory phonetic contexts for the spirantization of the target sounds. The spirantization rule was applied in some instances (for example, the word "dedo" was correctly produced as [deðo]) but not in others (for example, the word "dedo" was produced as *[dedo]). A total of 60 different stimuli were selected: 30 containing the correct fricative sound [ß, ð, γ], 10 per each segment, and 30 containing the incorrect voiced stop sound [b, d, g], 10 per each segment (Appendix A). The stimuli were tape-recorded in a soundproof booth. Before the experiment started, the intelligibility of the stimuli (provided by a female native speaker of Spanish) was assessed by four native speakers of Spanish: one female and three males. Listeners indicated whether the word they heard was correctly or incorrectly. Identification accuracy was 100% for all stimuli.

The stimuli for the production test consisted of an 80-word Spanish written paragraph containing 38 instances of the target sounds: 13 words for the sound [ß], 18 for [ð], and seven for [γ] (Appendix B).

Procedure

The experiment consisted of a perception test and a production test. Both tests were conducted in a language lab, where the participants were tested in a soundproof room. For the perception test, participants were presented a Spanish word on a screen. Once they saw the word, they clicked on it to hear the word pronounced. After listening to the word, they had to identify it as "correct" or "incorrect" on an answer sheet. Prior to the perception test, participants received a training session consisting of five test items to familiarize them with the tasks. For the production test, the participants recorded a paragraph containing several instances of the target sounds.

Perception test

Participants from the three proficiency levels (low, intermediate and advanced) took the perception test, in which they were presented with 60 randomized stimuli, with an inter-trial-interval of three seconds. The listeners were told to respond after each stimulus. They were encouraged to guess if unsure. The perception test lasted about 10 min, with one listener tested per day. All listeners were tested within a two-week period.

Production test

Immediately after the perception test, the participants took the production test. Before recording the paragraph, participants were given a minute to read the paragraph to familiarize themselves with it. The participants had only one chance to read the paragraph. The production test lasted about a minute, with one listener tested per day. All speakers were tested within a two-week period.

Data analysis

For the perception test, errors in identifying the tokens as "correct" or "incorrect" were counted and percentages of errors were calculated. For the production test, all words containing the target sounds were physically cut from each recording using Audacity. An initial aural analysis by the researchers (both native speakers of Spanish) was acoustically

	Low	Intermediate	Advanced

confirmed using the speech analysis program Praat (2010), and percentages of errors were calculated.

RESULTS

Perception

For the perception test, the number of potential errors, number of actual errors, and percentages of errors in perception were obtained for each group and target segment (Table 2).

Table 2. *Number of Potential Errors, Number of Actual Errors, and Percentages of Errors in Perception*

	# of potential errors	# number of errors	%	# of potential errors	# number of errors	%	# of potential errors	# number of errors	%
Bilabial	120	48	40	120	40	33	120	34	28
Dental	120	48	40	120	26	22	120	29	24
Velar	120	47	39	120	23	19	120	29	24

A comparison across proficiency levels per each target segment shows that bilabial errors decrease with higher level of proficiency (Low: 40% of errors, Intermediate: 33% of errors, Advanced: 28% of errors). Dental and velar errors show a very similar pattern across levels, with percentage of errors decreasing from Low to Intermediate (Dentals: Low: 40% of errors, Intermediate: 22% of errors; Velars: Low: 39% of errors, Intermediate: 19% of errors,), but staying stable at the Advanced level (Dentals: 24% of errors; Velars: 24% of errors). Results also indicate that bilabial sounds pose the most difficulties in perception for Intermediate and Advanced levels of proficiency (Low: 40% of errors, Intermediate: 33% of errors, Advanced: 28% of errors). See Figure 1 for a visual display of these results.

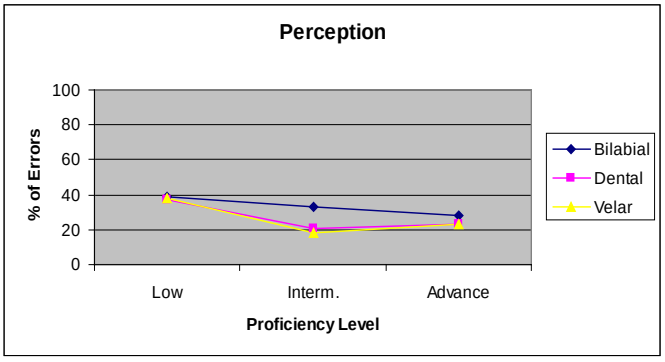


Figure 1. Perception results

A comparison within proficiency levels show that all three target sounds present the same level of difficulty for low proficiency participants (40%, 40% and 39% of errors respectively). For intermediate proficiency participants, bilabials (33% of errors) are the most difficult to perceive followed by dentals (22% of errors) and velars (19% of errors). For advanced proficiency participants, all three articulations present a similar level of difficulty, with bilabials being slightly more difficult than the others (28%, 24% and 24% of errors respectively).

Production

For the production test, the number of potential errors, number of actual errors, and percentages of errors were obtained for each group and target segment (Table 3).

Table 3. *Number of Potential Errors, Number of Actual Errors, and Percentages of Errors in Production*

	Low			Intermediate			Advanced		
	# of potential errors	# number of errors	%	# of potential errors	# number of errors	%	# of potential errors	# number of errors	%
Bilabial	78	37	48	78	21	27	78	20	26
Dental	126	96	76	126	58	46	126	53	42
Velar	42	15	36	42	14	34	42	14	33

A comparison across proficiency levels per each target segment shows that bilabial and dental errors decrease from Low to Intermediate levels of proficiency (Bilabials: 48% and 27% of errors respectively; Dental: 76% and 46% of errors respectively) and stays more or less the same from intermediate to advanced (Bilabial: 27% and 26%; Dentals: 46% and 42% respectively). Velar errors are similar for all levels of proficiency (36%, 34% and 33% of errors). Results also indicate that dental sounds pose the most difficulties in production for all three levels of proficiency (Figure 2).

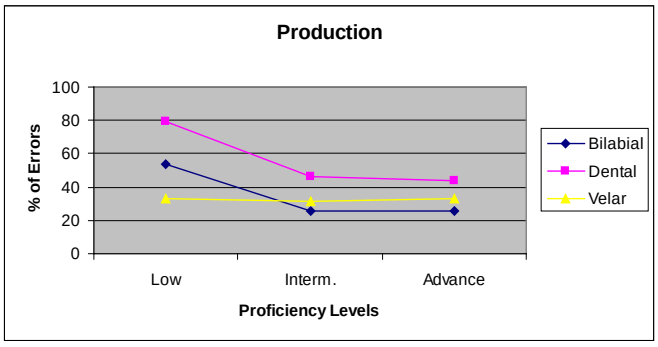


Figure 2. Production results

A comparison within proficiency levels shows that for Low proficiency participants, dentals are the most difficult (76% of errors), then bilabials (48% of errors), and velars are the easiest (36% of errors). For Intermediate proficiency participants, dentals are also the most difficult (46% of errors); whereas bilabials (27% of errors) and velars (34% of errors) show relatively similar levels of difficulty. For advanced proficiency participants, dentals (42% of errors) are the most difficult, then velars (33% of errors), and bilabials (26% of errors) are the easiest.

Perception versus Production

The percentages of errors from each test were compared for each group and target segment (Table 4).

Table 4. *Percentage of Errors in Production and Perception (%)*

	Low		Intermediate		Advanced	
	Perception	Production	Perception	Production	Perception	Production
Bilabial	40	48	33	27	28	26
Dental	40	76	22	46	24	42
Velar	39	36	19	34	24	33

Results show that there is a tendency for bilabials and velars to be perceived and produced with similar degrees of accuracy in all three levels of proficiency. It is in the dental place of articulation where perception and production differs the most, with production being more difficult than perception in all three levels of proficiency (Figures 3, 4, and 5).

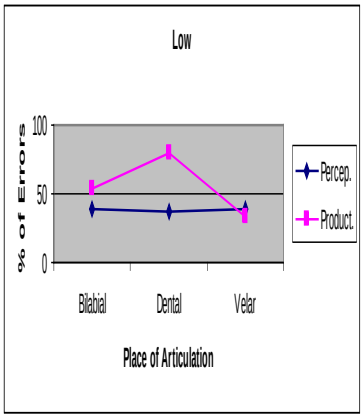


Figure 3. Low Proficiency Level

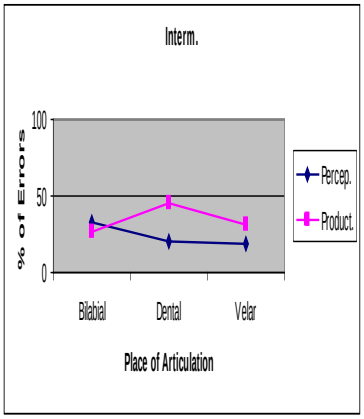


Figure 4. Intermediate Proficiency Level

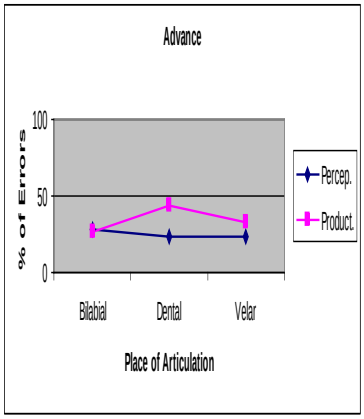


Figure 5. Advanced Proficiency level

DISCUSSION

Each of the research questions will be discussed in turn.

R.Q. #1: At what levels of proficiency are L2 Spanish learners able to recognize whether the spirantization of Spanish voiced stops rule has been applied, as indicated by the results of the perception test?

Participants at the Intermediate level of proficiency started to recognize the spirantization rule in dentals and velars, but still had some difficulties perceiving spirantized bilabials. Those at the Advanced level showed no difference from the Intermediate ones in perceiving spirantized dentals and velars, but did better in the perception of spirantized bilabials. So the answer to question 1 is that L2 learners of Spanish are able to recognize the spirantization of dentals and velars at the Intermediate Level of proficiency, and is not until the Advanced level that the spirantization of bilabials is shows the same level of recognition.

As for the reason why learners at the Intermediate level have trouble perceiving the spirantization of [b], we might hypothesize that the English phoneme [v] interfered with the perception of [ß], since "v" is a plausible spelling in Spanish of the phoneme /b/. When participants read words containing a "v" (such as "ave" or "cava,") during the perception part of the study, they might have expected it to be pronounced as [v], and when they heard [ß] instead, they labeled the word as "incorrect." On the other hand, when they saw a Spanish word spelled with a "b" and then heard the unfamiliar [ß], they might have taken it as a [v], and therefore labeled the word as "incorrect." This hypothesis falls along the lines of Zampini's (1994) when she also speculates that orthography may have played a role in the pronunciation of /b/ by the native English speakers in her study.

R.Q. #2: At what levels of proficiency are L2 Spanish learners able to apply the spirantization of Spanish voiced stops rule, as indicated by the results of the production test?

Participants at the Low Level of proficiency already produced spirantized velars as well as Intermediate and Advanced participants, but had more problems producing bilabials and especially dentals. It is also interesting to see that there is no difference in the production of spirantized voiced stops between the Intermediate and Advanced levels of proficiency. So the answer to question 2 is that L2 Spanish learners are able to apply the spirantization of Spanish voiced stops rule at the Low level for velars, and at the Intermediate Level for bilabials and dentals.

The difficulty that English-speaking learners of L2 Spanish have in acquiring Spanish dental sounds has been repeatedly reported in the literature (Bowen & Stockwell, 1957; González-Bueno, 1995, 1997, 2002; Macken & Barton, 1979). Macken and Barton (1979) specifically say that "Labial stops are most likely to be spirantized and dental stops are least susceptible to spirantization" (p. 447). There are two main reasons why the spirantization of Spanish voiced stops is particularly difficult for native speakers of English. One is the fact that the sounds [d] has an alveolar articulation in English, unlike the dental articulation in Spanish. The mere anatomical aspect of this alveolar articulation prevents spirantization of this sound. Rather, the weakening of English alveolar sound [d] manifests itself in the process of "flapping" the sound [d] rather than spirantizing it.

R.Q. #3: What is the relationship between L2 Spanish learners' perception and production results, as indicated by both tests?

The results show that there is a tendency for bilabial and velars to be perceived and produced similarly at all three levels of proficiency and that it is in the dental place of articulation where perception and production differ the most, with production being more difficult than perception for all three levels of proficiency. The reasons for this have already been addressed when discussing the answer to and conclusions from research question #2, that is, different place of articulation and the application of the English flapping process.

We initially intended to analyze the effect of a Phonetics course on the perception and production of the target sounds by Advanced proficiency level participants. However, given the small number of participants at this level of proficiency, the results of this analysis were rendered non-significant. Nevertheless, these results are included as anecdotal evidence to support previous studies on the effect of formal instruction on pronunciation.

Production and perception results, in the form of percentages of errors, from the advanced participants (n=6), were separated in two groups (Table 5): Those who took a Phonetics Course (n=4) and those who did not (n=2).

Table 5. *Effect of Phonetics Course*

Phonetics Course	Perception (% of errors)	Production (% of errors)
Yes (n. 4)	28	30
No (n. 2)	18	48

These results seem to suggest that formal instruction in Spanish Phonetics shows an effect on production but not on perception. Participants who took a Phonetics course were perceived spirantization somewhat worse (28% of errors) than those who did not (18% of errors). However, those who took a Phonetics course produced spirantized sounds better than those who did not (Figure 6).

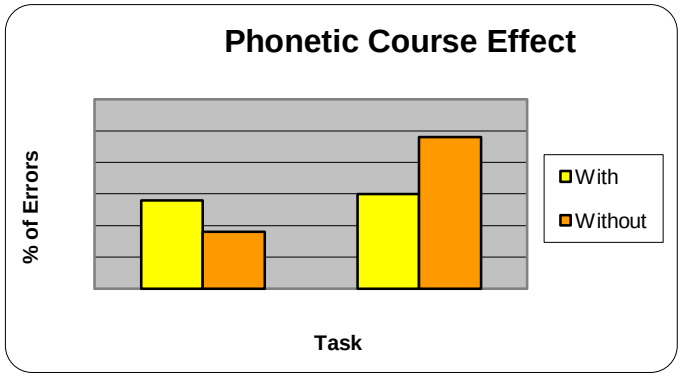


Figure 6. Phonetics Course Effect Results

Lord (2010) also found a positive effect of a Phonetics course on the production of the voiced stop's spirantized allophones. However, a Phonetics course did not show an effect in the perception part of the present study (participants who took a Phonetics course perceived spirantized sounds worse than those who did not).

This might sound counter-intuitive, since in general, perception develops earlier than production, as reported by previous studies (Bradlow, Pisoni, Akahane, & Tohkura, 1997; Rochet, 1995; Wang, Jongman, & Sereno, 2003.) Aside from the effect of the formal instruction in pronunciation, we can observe that although learners at the Advanced level made fewer errors in production (M = 33.7% of errors) than the Intermediate level (M = 35.7% of errors) this difference is not as dramatic as the one between the Intermediate level and the general production of the Low level (M = 53.3% of errors). This leads us to speculate that general exposure to the target language seems to have a definite positive effect on the pronunciation of the fricative allophones of Spanish voiced stops. Furthermore, if the Phonetics course has the effect of moving advanced level students slightly up in the scale of accuracy, the effect might be greater if this Phonetics instruction were available to lower level learners.

In view of these results, we can conclude that, at least for the participants in this study, level of proficiency has a direct effect on the level of awareness of the Spanish spirantization rule. Learners at the Intermediate level start showing awareness of the Spanish spirantization rule in perception and production: In perception: for velars and dentals, but not for bilabials, and in production, in all three places of articulation. When comparing how the spirantization rule is recognized with its application in production, there is a relationship between the perception and production of the fricative allophones of bilabials and velars, but not of dentals. Both bilabials and velars seem to present similar levels of difficulty in both perception and production. For dentals, on the other hand, the results show that they are always more difficult to produce than to perceive, and this happens in all three levels of proficiency.

An additional conclusion that can be drawn from this study is that of the order of appearance of the target sounds in the interlanguage of learners. The rank order of fricative allophones in González-Bueno's (1995) study resembles the results obtained by Zampini (1994). Learners produced [ɣ] most frequently, [ß] was the second most

frequently pronounced fricative, and [ð] was produced the least. In the present study, the observed order of acquisition in production is also similar only for learners at the Low Level proficiency: low level participants made fewer errors in the production of [ɣ], followed by [ß], and they made the most errors in the production of [ð]. However, for learners at the Intermediate and Advanced levels, the order in terms of number of errors is slightly different: they made fewer errors in [ß], followed by [ɣ], and they made the most errors in the production of [ð]. In other words, [ß] and [ɣ] exchanged orders, but [ð] was still the most difficult one (Table 6).

Table 6. *Rank Order in Production*

Rank Order	Zampini (1994)	González-Bueno (1995)	This study		
			Low	Intermediate	Advanced
1st	[ɣ]	[ɣ]	[ɣ]	[ß]	[ß]
2nd	[ß]	[ß]	[ß]	[ɣ]	[ɣ]
3rd	[ð]	[ð]	[ð]	[ð]	[ð]

This comparison reinforces this study's conclusion of [ð] being the Spanish sound that is more difficult to produce of the three fricative allophones. In terms of perception, we saw that it was [ß] the most difficult to perceive, but we cannot compare this with previous studies, since they did not address perception of [ß, ð, ɣ], only production.

LIMITATIONS AND RESEARCH RECOMMENDATIONS

The results of the present study cannot be taken as definitive given the two following drawbacks:

1. Low number of participants. Recruiting participants was one of the most challenging tasks in this investigation. Despite personal visits to different Spanish classes, display of flyers, electronic messages to Spanish students inviting them to participate, recruitment was extremely low. One obstacle was that the Spanish Department would not give extra credit for students' participation in the research (given the tight grading system already in place.) We even went as far as to offer $5 in order to recruit enough Spanish students so we could have a reasonable number of participants. Similar studies should be carried out with larger numbers of participants. A greater population sample will allow for greater confidence in the results and findings.

2. Low validity of the approach to determine participants' language proficiency level:

The participants' levels of language proficiency were already determined by the course in which they had been placed by the Spanish Department. For convenience reasons, we did not conduct any additional test that might have provided a more valid assignment of participants' proficiency levels. Future studies should include in their methodological design a more valid assessment procedure to more accurately determine participant's levels of proficiency.

3. Undetermined previous instruction: It was assumed that, aside from the Phonetics course taken by some of the participants at the Advanced level, no formal instruction on Phonetics was provided. It could be that some participants might have received this kind of instruction during their Spanish studies, which was not controlled for in this research.

Notwithstanding the methodological drawbacks, the contributions of this study should not be disregarded, considering the implications outlined below.

Pedagogical Implications

Given that dental sounds pose the most difficulty in production for all three levels of proficiency, it can be assumed that one cause of this difficulty may be the different place of articulation of this sound in Spanish (dental) and English (alveolar). Therefore, the first goal of a training program to help with the pronunciation of the voiced dental allophone of [d] is to make learners aware of the different place of articulation of Spanish dentals. An accurate anatomical description of the oral cavity to show the location of both the alveolar ridge (English pronunciation) and the back of the frontal teeth (Spanish pronunciation) should help achieve this goal (Figures 7 and 8).

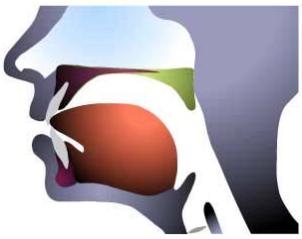


Figure 7. Spanish Dental articulation of [d]

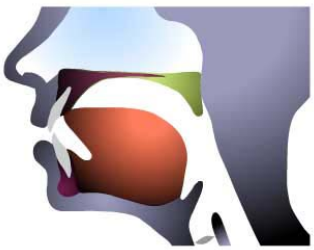


Figure 8. English Alveolar articulation of [d][1]

Learners should also made be visually aware of the articulation of the interdental allophone of the voiced dental Spanish sound, which is only slightly different from the articulation of their familiar English [ð] used to pronounce the grapheme "th" as in "they" (Figures 9 and 10).

[1] Images taken from *The Sounds of Spoken Language* (www.uio.edu/~acadtech/Phoneticss/)

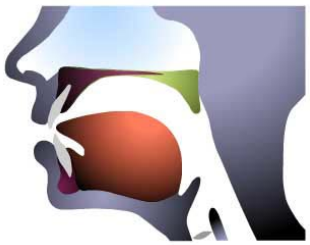


Figure 9. Spanish Interdental articulation of [ð]

Figure 10. English Interdental articulation of [ð]

In addition, controlled practice of the target sound [ð], both in perception but most importantly in production, should be provided. The very design of the perception test used in this study could serve as a model to design these perception exercises. For example, learners should first be asked to determine whether the intervocalic [d] in the Spanish "nada," for instance, has been correctly pronounced as [ð] or incorrectly as [d]. Then, attention should be paid to the accurate pronunciation of [ð] by having learners pronounce phrases such as "un dato," "el dato" where both instances of "d" are pronounced as [d], and "¿qué dato?" in which "d" is pronounced as [ð] (Table 7).

Table 7. *Examples of phrases with instances of [ð]*

	[d]		[ð]
Dedo	un dedo	el dedo	mi dedo
Duna	gran duna	mil dunas	la duna
Dama	gran dama	tal dama	una dama
Dote	sin dote	tal dote	la dote

Similarly, learners should be made aware of the Spanish articulation of intervocalic [b] by providing them with an accurate anatomical description of the oral cavity to show the articulation of the fricative allophone [ß] (Figures 11 and 12).

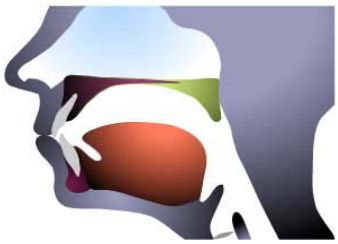


Figure 11. Spanish stop articulation of [b]

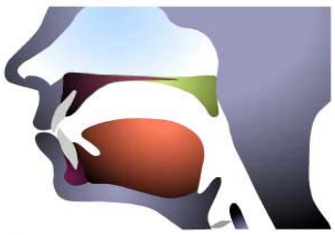


Figure 12. Spanish fricative articulation of [ß]

The results of this study also indicated that the sound [ß] poses the most difficulty in perception for all three levels of proficiency. Therefore, controlled practice of the target sound [ß] in perception, should be provided. As suggested above for the training of [ð]), perception exercises could follow the model of the perception test used in this study. For example, learners should first be asked to determine whether the intervocalic [b] in the Spanish word "bota," for instance, has been correctly pronounced as [ß] or incorrectly as [b]. Then, attention should be paid to the accurate pronunciation of [ß] by having learners pronounce phrases such as "un beso," in which "b" is pronounced as [b], and "el beso" where "b" is pronounced as [ß] (Table 8).

Table 8**.** *Examples of phrases with instances of [ß]*

[b]		[ß]
burro	un burro	mi burro
bota	con bota	la bota
beso	un beso	tu beso
balón	un balón	su balón

Special attention should be given to the fact that the grapheme "v" is pronounced in identical way as "b," and therefore the spirantization rule applies equally to it. To that end, learners should practice with phrases such as "un vaso" in which "b" is pronounced as [b] and "el vaso" where"b" pronounced as [ß] (Table 9):

Table 9. *Examples of phrases with instances of grapheme "v" pronounced as [b] and [ß]*

[b]		[ß]
veneno	con veneno	el veneno
visado	sin visado	mi visado
vuelta	gran vuelta	una vuelta
valor	con valor	su valor

It goes without saying that instructors should implement the controlled practice described above in a contextualized and meaningful way, the design of which is beyond the scope of this study.

With the present study, we hope to have contributed to the line of research focusing on the acquisition (both perception and production) of Spanish fricative allophones of voiced sounds. There is a growing global demand for accurate communication in a foreign language, and pronunciation plays an important role in accurate oral performance. A strong foreign accent can interfere with communication, and the mispronunciations of the sounds addressed in this study greatly contribute to a foreign accent. Therefore, we urge Spanish teachers to adopt the pedagogical strategies presented here to help students overcome the difficulties of these sounds, and become efficient communicators in the target language. Furthermore, we recommend that this pronunciation training be available to learners at all levels of proficiency.

REFERENCES

Bowen, J.D. & Stockwell, R.P. (1957). Orthography and respelling in teaching Spanish. *Hispania, 40*, 200-205.

Bradlow, A., Pisoni, D., Akahane, R. & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: Some effects of perceptual learning in speech production. *Journal of the Acoustical Society of America 101*, 2299-2310.

Dalbor, J. B. (1997). *Spanish pronunciation, theory and practice*. Orlando, FL: Harcourt Brace & Company.

Díaz-Campos, M. (2004). Context of learning of Spanish second language phonology. *Studies in Second Language Acquisition, 26,* 249-273.

Elliott, A.R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania, 80,* 96-108.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning 27,* 315-330.

González-Bueno, M. (1995). Adquisición de los alófonos fricativos de las oclusivas sonoras españolas por aprendices de español como segunda lengua. *Estudios de Lingüística Aplicada, 13,* 64-79.

González-Bueno, M. (1997). The Effects of formal instruction on the acquisition of Spanish phonology. *Contemporary perspectives on the acquisition of Spanish.* Ed. by W. R. Glass and A.T. Pérez-Leroux. Cascadilla Press: Somerville, Massachusetts.

González-Bueno, M. (2002). Dental versus alveolar articulation of L2 Spanish stops as perceived by native speakers of Malayalam. In Shoshuke Haraguchi, Bohumil Palek, and Osamu Fujimura (Eds.) *Proceedings of "Linguistics and Phonetics 2002"* (LP2002,) Charles University Press and Meikai University, Japan.

Hualde, J. J. (2005). *The sounds of Spanish*. Cambridge: Cambridge University Press.

Jakobson, R. (1941). *Child language: aphasia and phonological universals*. The Hague: Mouton.

Krashen, S., & Terrl, T. D (1983). *The natural approach*. Pergamon.

Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. *Proceedings of the XIII International Congress of Phonetics Science* (pp. 92-99). Stockholm University.

Lord, G. (2010). The combined effects of instruction and immersion on second language pronunciation. *Foreign Language Annals, 43, 4,* 488-503.

Macken, M.A. & Barton, D. (1979). The acquisition of voicing contrast in Spanish: a phonetics and phonological study of word-initial stop consonants. *Child Language, 7,* 433-58.

Rochet, B.L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Baltimore, MD: York Press.

Schwegler, A., & Kempff, J. (2007). *Fonética y fonología españolas: teoría y práctica* (3rd Ed.) New York: John Wiley & Sons.

Wang, Y., Jongman, A. & Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113 (2),*1033-1043.

Zampini, M. (1994). The role of native language transfer and task formality in the acquisition of Spanish spirantization. *Hispania, 77,* 470-481.

Appendix A

STIMULI LIST

	[ß]
1.	Haba
2.	Iba
3.	Bebe
4.	ave
5.	Sabe
6.	baba
7.	Boba
8.	cava
9.	lobo
10.	hubo

	[ð]
1.	Dedo
2.	Nada
3.	Oda
4.	Codo
5.	Lodo
6.	Lado
7.	Mide
8.	Pide
9.	hada
10.	mudo

	[ɣ]
1.	Hago
2.	Mago
3.	Lego
4.	miga
5.	fuga
6.	higo
7.	siga
8.	logo
9.	daga
10.	llego

Appendix B

READING PASSAGE

La cueva del lobo

El mago llegó a la cueva todo cubierto de lodo y baba, asustando a las aves posadas en una higuera que había al lado de la entrada. El lobo, antes de darse a la fuga, le había mordido en el codo y en los dedos, y ahora apenas podía sostener la daga con la que intentó defenderse. Llegó hasta el lago para lavarse y beber un poco. No se oía nada, era como si el bosque se hubiera quedado mudo.

Ruellot, V. (2011). Computer-assisted pronunciation learning of French /u/ and /y/ at the intermediate level. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2010. (pp. 199-213), Ames, IA: Iowa State University.

COMPUTER-ASSISTED PRONUNCIATION LEARNING OF FRENCH /u/ AND /y/ AT THE INTERMEDIATE LEVEL

Viviane Ruellot, Western Michigan University

This study explores the impact of visual feedback derived from a speech analysis program designed to improve the pronunciation of French /u/ (as in *tout* [tu], "all") and /y/ (as in *tu* [ty], "you") in the speech of intermediate-level adult learners of French with L1 English. As /y/ is absent from the English phoneme repertoire and represents a new phone for these learners, successful distinction between French /u/ and /y/ is contingent upon experience with the language (Flege, 1987). Visual representation of articulators, with a focus on tongue position (back for /u/ and front for /y/), may help learners create distinct phonemic categories for these sounds. Students (*n*=14) in a third-year French phonetics course recorded their pronunciation of French monosyllabic words featuring /u/ and /y/. Participants in the audio-visual condition (*n*=7) received visual information about the formant trajectories of /u/ and /y/ and instruction as to their correlation with degree of mouth aperture (F1), tongue backing (F2), and lip rounding (F3). The recordings were assessed by native French speakers. Results indicate that the presence of visual feedback did not significantly improve pronunciation. The relation between the efficacy of visual pronunciation feedback, time on task, and perception skill development are discussed.

INTRODUCTION

Alternately considered marginal and as important as grammar and vocabulary, the status of pronunciation in the history of the foreign language curriculum has widely fluctuated. In the early 21st century, this skill has regained importance in instruction (Jenkins, 2004). While the efficacy of explicit pronunciation instruction has been tested (with positive results, e.g., Couper, 2003; Derwing, Munro, & Wiebe, 1998; Derwing & Rossiter, 2003), research on pedagogical tools, and more particularly instructional technology, has yielded mixed results. After a review of the relevant research in this area, this paper reports on two experiments testing the efficacy of the visual feedback generated by the speech analysis program WaveSurfer (Sjölander & Beskow, 2000-2005).

BACKGROUND

The current presentation and practice of pronunciation in textbooks and workbooks remains underdeveloped for at least two reasons. First, pronunciation presentation is commonly limited to verbal indications, unaccompanied by the support of visual elements (e.g., representations of articulators) which may help learners put into practice the information given and allow them to encode and retrieve the information through more than one cognitive channel (i.e., aural/oral and visual) (Paivio, 1971; 1991). Second, the opportunities for feedback on the learner's production are generally limited in both paper and electronic versions of instructional materials, and they rarely meet the

criteria for effective feedback, which is "comprehensible, [does] not rely solely on the learner's own perception, [allows] verification of response correctness, [pinpoints] specific errors and possibly [suggests] a remedy" (Neri, Cucchiarini, & Strik, 2002, 1210).

With the development of automatic speech recognition technology, commercial pronunciation software programs have begun to address the above-mentioned issues. However, improvements are still needed, particularly regarding the adequacy of the linguistic and pedagogical content of the programs (Derwing & Munro, 2005; Levis, 2007). Still, one feature, the visual representation of speech, which is available in these programs but also in software designed for phonetics and phonology research, has caught the eye of both pedagogues and researchers who have studied the extent of the contribution of spectrograms, waveforms and pitch trackers to the development of pronunciation skills (Levis, 2007; O'Brien, 2006). A pioneer in the field, de Bot (1983) demonstrated the contribution of visual feedback to the improvement of intonation for Dutch learners of English. Hardison (2004) investigated the effectiveness of visual feedback on the learning of French prosody and segments and found significant improvement at both levels, as well as heightened awareness of prosodic and segmental features, improved listening comprehension skills, and increased overall confidence about pronunciation.

Experiments focusing on pitch contours and intonational patterns have more often yielded success and significant pronunciation improvement than have studies concentrating on segments alone (Levis, 2007). Among the oft-cited limitations of the spectrogram as a source of visual feedback on segments is the assumption that the acoustic representation of the learner's utterance should closely match that of the target model (Neri et al., 2002). This assumption does not seem realistic due to the considerable anatomical variations that exist among speakers and which largely affect the acoustic representation of their utterances. Another limitation of spectrographic displays is related to the complex technical information they provide (Neri et al., 2002). Learners are rarely familiar with such displays and their associated technical terms (e.g., "formants," "frequency," "hertz," etc.), and generally need the help of a trained informant to interpret them. Finally, one of the reasons behind the greater success of experiments focusing on pitch and intonation may be related to the iconic nature of the visual representation in which the "rising, falling, and level lines (…) usually [correspond] to rises, levels, and falls in a speaker's voice pitch" (Levis, 2007), making the visual feedback a more direct and readily interpretable one.

These issues were addressed in an experiment (Ruellot, 2007) in which beginner American students of French practiced their pronunciation of words featuring the vowels /a/, /i/, /u/ (as in *tout* 'all'), and /y/ (as in *tu* 'you'). The subjects in the experimental group studied spectrograms whose degree of iconicity was enhanced through color-coding of formant trajectories correlating with articulators and which excluded technical terms (Guilloteau, 1998). The subjects in the experimental group significantly improved their pronunciation of words featuring /u/ presented in isolation and in sentences. However, none of the subjects significantly improved their pronunciation of items in /y/. Although the presence of visual feedback was met with general enthusiasm, a majority of subjects

expressed the need for both additional time and interactive help from an instructor to interpret the visual displays.

The present study addressed these issues and the experimental design in the original study (i.e., Ruellot, 2007) was modified to provide subjects with additional practice time and repeated interaction with an instructor for help on display interpretation, as well as tips on how to improve pronunciation (Engwall, Wik, Beskow, & Granström, 2004). In addition, the sounds at study were limited to French /u/ and /y/, which are typically challenging for English native speakers, as /y/ does not exist in their native repertoire. Principles of Dispersion Theory (Liljencrants & Lindblom, 1972; Lindblom, 1986a; 1986b; 1990; Lindblom & Maddieson, 1988), maximal vowel dispersion (Maddieson, 1984), and predictions from the Speech Learning Model (Flege, 1987) point to the assimilation of /y/ in /u/ during the first stages of acquisition, and the development of /y/ into a distinct phonemic category as experience with the language increases. In order to favor such development, the experimental design in the present study also included a condition in which visual feedback of /u/ and /y/ was presented side-by-side, allowing subjects to concentrate on tongue position (i.e, the main contrast between these two sounds). It was assumed that such focus on a specific articulator would allow subjects to identify their mistakes (Engwall et al., 2004; Pennington & Ellis, 2000) and favor its correction.

The following two hypotheses were tested in two separate experiments in the present study:

Additional practice time and interactive explanations and feedback from an instructor will significantly help subjects receiving visual feedback improve their pronunciation of items featuring /u/ and /y/ (Experiment 1).

Visual feedback of /u/ and /y/ presented side-by-side and studied together will help develop awareness of tongue position, which will result in significantly higher scores than when feedback for /u/ and /y/ is presented and studied separately (Experiment 2).

EXPERIMENT 1: STUDY OF VOWELS IN ISOLATION

Experimental design

Participants

The students, enrolled in a third-year French phonetics course at an American university, were randomly assigned to one of two groups: the Audio feedback condition (n = 7), and the Audio-visual feedback condition (n = 7).

Instruments

Training for all subjects included a reminder of the articulatory contrasts between French /u/ and English /u/, and between French /u/ and /y/. While the subjects in the Audio condition used basic audio recording software[ix], the subjects in the Audio-visual condition used WaveSurfer (Sjölander, & Beskow, 2000-2005), the same freely available

spectrographic analysis program used in the original study. The subjects in the Audio condition were asked to step outside of the classroom for 6 minutes, during which the subjects in the Audio-visual condition were instructed about the spectrographic displays they would study.

These displays were simplified in the same manner as in the original study: technical terms were not addressed and, in order to increase the iconic aspect of the displays, the first three formants of each vowel were color-coded and labeled according to their corresponding articulators.[x] The instructions for the interpretation of the visual displays were also replicated from the original study: subjects were informed that the three lines move vertically according to 1) how open the mouth is, 2) the position (front or back) of the tongue, and 3) the degree to which lips are rounded (CALLIOPE, 1989). The subjects' attention was particularly drawn to the relative distances between the native speaker's lines,[xi] and the subjects' task was to approximate these distances, not the actual 'heights' (i.e., hertz values) of the lines, following recommendations as to how to modify the position of articulators also included in the instructions. In the original study, these instructions and recommendations were provided exclusively in writing. In the present study, they were additionally presented by the instructor orally, and revisited during individual subject-instructor interactions throughout the experiment.

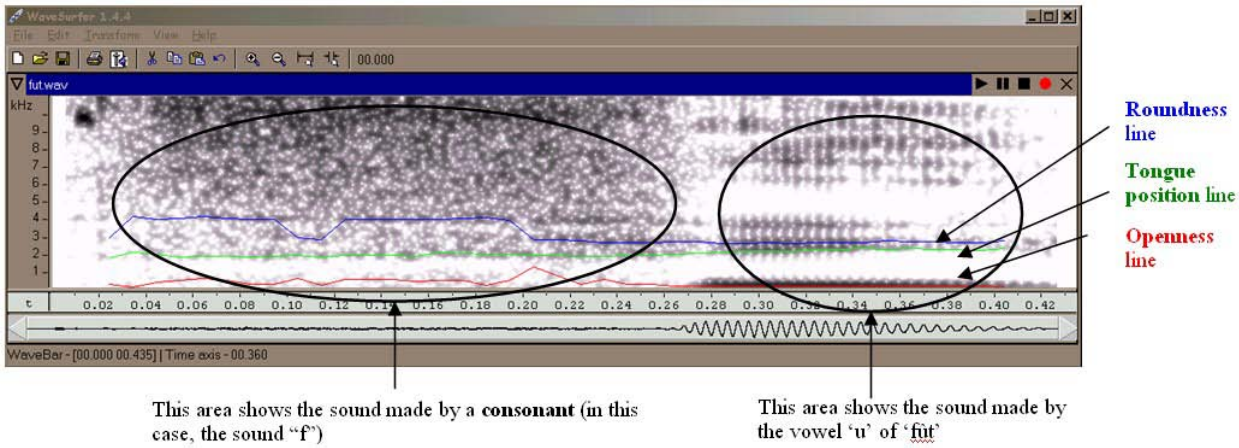


Figure 1. Model Spectrogram of fût 'barrel' uttered by a Female Native Speaker Model (Excerpt from the Instructions provided to the Subjects in the Audio-visual Condition)

Stimuli

All the subjects studied /u/ and /y/ through 10 monosyllabic words, which they recorded, along with sentences featuring them (see Table 1) in order to verify transfer of pronunciation improvement to a contextualized environment.

Table 1. *Stimuli*

Words (rehearsed)[xii]	/u/	*fou* 'crazy', *loup* 'wolf', *mou* 'soft', *nous* 'we', *sous* 'under'
	/y/	*fût* 'barrel', *lu* 'read' (past participle), *mue* 'break' (related to voice), *nus* 'naked', *sue* 'sweat'
Sentences (not rehearsed)	/u/	*Eric est fou de son chien Léon.* 'Eric is crazy about his dog Leon.'
		Le loup est un animal sauvage. 'The wolf is a wild animal.'
		Le canapé d'Anne est mou mais confortable. 'Anne's sofa is soft but comfortable.'
		Ce soir, nous dansons le rock. 'We're dancing to rock dances music.'
		Les chaussures vertes sont sous la table. 'The green shoes are under the table.'
	/y/	*Le vin reste dans un fût pendant cinq ans.* 'Wine is kept in a barrel for five years.'
		Jacques a lu un très bon livre sur les Incas. 'Jacques read a very good book about the Incas.'
		La voix des garçons mue durant l'adolescence. 'Boys' voice breaks during adolescence.'
		Boticelli aime peindre des nus. 'Boticelli likes to paint nudes.'
		Il a chaud et il sue beaucoup. 'He's hot and he sweats a lot.'

Treatment and Tests

Subjects participated in two 50 minute-long pronunciation sessions devoted to the study of /u/ (session 1) and /y/ (session 2). They listened to native speaker model pronunciations of each target word, practiced and listened to their own pronunciation as often as they wished for three minutes, after which they recorded their final version of the target word. Each session ended with the recording of unrehearsed sentences featuring the practiced words.

In the original study, time allocated to the Audio-visual condition for the study and practice of the first word in each target vowel series amounted to 12 minutes. During this time, subjects familiarized themselves with the visual displays and their accompanying written explanations. In the present study, this time was extended to 18 minutes (in addition to the 6 minutes of initial introduction to spectrographic displays mentioned above), during which a trained instructor interacted with subjects, guiding their

interpretation of the visual displays and their reflection as to how to modify their articulators to approximate target pronunciations. The instructor continued to provide individual help during the rest of the sessions. The subjects in the Audio condition were also assigned a trained instructor who guided them, throughout the experiment, in their reflection on how to modify their articulators for successful pronunciation. All subjects recorded their pronunciation of the isolated words and sentences three times: before, immediately after and, in order to test for long-term improvement, seven days after treatment.

Rating Procedure

The 840 recordings obtained were submitted for assessment to three French native speakers who rated two types of recordings: 1) words pronounced in isolation and 2) the same words embedded in sentences. The latter were extracted from their sentence context so as to minimize the impact that the non-target words and sounds may have had on the ratings. The raters used a 5-point Likert-type scale (Bongaerts, Planken, & Schils, 1995) ranging from 1 – "very strong foreign accent; the comprehension of the word is very difficult; this speaker is undeniably not a native speaker of French", to 5 – "no trace of a foreign accent; the comprehension of the word is very easy; this speaker is undeniably a native speaker of French."

RESULTS

The first hypothesis posits that more time on task and interactive explanations with – and feedback from – the instructor will help learners who receive visual feedback improve their pronunciation of /u/ and /y/.

- Short-term improvement

Although all of the subjects in the experiment improved their pronunciation of words in /u/ and /y/, as well as target words in /y/ appearing in sentences (i.e., extracted words) immediately after treatment (Table 2), improvement in the Audio-visual condition was not significantly different than in the Audio condition. This can be concluded from the results of unpaired t-tests run on mean ratings obtained immediately after treatment for words and extracted words in /u/ and /y/ with condition as a between-subjects factor displayed in Table 3.

Table 2. *Mean Ratings and Mean Difference in Mean Ratings from Pre-test to Post-test for Isolated Words and Extracted Words (both Conditions combined)*

Vowel	Recording	Post-test mean	Pre-test mean	Mean difference	*t*	*df*	*p*
u	Words	3.524	2.941	0.583	3.477	13	0.004**
y	Words	3.581	3.357	0.224	1.849	13	0.001**
y	X-words	3.783	3.010	0.773	4.181	13	0.019**
u	X-words	3.636	3.092	0.543	2.672	13	0.087

**p <.01; 'X-words' refers to extracted words

Table 3. *Mean Ratings and Mean Difference in Mean Ratings at the Post-test level for Isolated Words and Extracted Words in the Audio and the Audio-visual Conditions*

Vowel	Recording	Condition	Mean	Mean difference	*t*	*df*	*p*
/u/	Words	Audio-visual	3.572	0.096	0.307	12	0.764
		Audio	3.476				
/y/	Words	Audio-visual	3.661	-0.244	-0.489	12	0.634
		Audio	3.905				
/u/	X-words	Audio-visual	3.629	0.095	0.301	12	0.769
		Audio	3.533				
/y/	X-words	Audio-visual	3.700	0.130	0.225	12	0.826
		Audio	3.571				

'X-words' refers to extracted words

- Long-term pronunciation improvement

As can be gathered from Table 4, long-term improvement for all of the subjects in the experiment was restricted to words in /y/. That is, the pronunciation improvement for words in /u/ and extracted words in /y/ obtained immediately after treatment was not maintained one week later.

Table 4. *Mean Ratings and Mean Difference in Mean Ratings from Pre-test to Delayed Post-test for Isolated Words and Extracted Words*

Vowel	Recording	Delayed Post-test mean	Pre-test mean	Mean difference	*t*	*df*	*p*
u	Words	3.009	2.941	0.068	0.519	13	0.613
y	Words	3.685	3.010	0.675	3.756	13	0.002**
y	X-words	3.367	3.092	0.275	1.505	13	0.156
u	X-words	3.228	3.357	-0.128	-0.860	13	0.405

**p <.01; 'X-words' refers to extracted words

However, the presence of visual feedback did not significantly contribute to long-term improvement of words in /y/, as indicated by the absence in Table 5 of a significant difference between conditions.

Table 5. *Mean Ratings and Mean Difference in Mean Ratings at the Delayed Post-test level for Isolated Words and Extracted Words in the Audio and the Audio-visual Conditions*

Vowel	Recording	Condition	Mean	Mean difference	*t*	*df*	*p*
/u/	Words	Audio-visual	2.960143	-0.097	-0.220	12	0.830
		Audio	3.057143				
/y/	Words	Audio-visual	3.645500	-0.080	-0.149	12	0.884
		Audio	3.725429				
/u/	X-words	Audio-visual	3.133143	-0.191	-0.653	12	0.526
		Audio	3.323714				
/y/	X-words	Audio-visual	3.390571	0.047	0.092	12	0.928
		Audio	3.343143				

'X-words' refers to extracted words

EXPERIMENT 2: STUDY OF VOWEL PRESENTED IN PAIRS

Experimental design

Experiment 2 started one week after Experiment 1 and lasted for one session. The participants were the same students enrolled in the French phonetics course who had served as subjects in Experiment 1. For the second experiment, they switched groups, and the subjects from the Audio-visual condition in Experiment 1 were now assigned to the Audio condition and exclusively worked with audio feedback, while the subjects from the Audio condition in Experiment 1 now made up the Audio-visual condition and additionally studied visual displays of /u/ and /y/ presented side-by-side. Initial training, instruments, stimuli, treatments, tests, and rating procedure were almost the same as in Experiment 1. The only difference is that Experiment 2 took place over only one session, during which /u/ and /y/ were not studied separately but simultaneously.

RESULTS

In order to test hypothesis 2 and compare the efficacy of the visual feedback on pronunciation of specific sounds when they are presented, studied and practiced separately vs. side-by-side, two series of unpaired t-tests were run. Each contrasted mean ratings from the Audio-visual condition in Experiment 1, i.e., when/u/ and /y/ were studied separately, with mean ratings by the subjects in the Audio-visual condition in Experiment 2, when /u/ and /y/ were presented side-by-side. The first series of unpaired t-tests contrasts performance in the short-term, i.e. immediately after treatment, while the second one compares long-term performance, i.e., one week after treatment.

Although the subjects who studied the visual display of /u/ and /y/ presented side-by-side significantly improved their pronunciation of words in /u/ immediately after treatment (paired $t(6) = 7.316$, $p = 0.0003$), their scores were not significantly higher than those of the subjects who studied the vowels separately, be it immediately after treatment (see Table 6) or one week later (see Table 7). In other words, studying the articulatory contrasts between /u/ and /y/ through visual displays presented side-by-side did not constitute an advantage over studying them separately.

Table 6. *Mean Ratings and Mean Difference in Mean Ratings at the Post-test Level for Isolated Words and Extracted Words in the Audio-visual Conditions in Experiment 1 (Audio-visual 1) and Experiment 2 (Audio-visual 2)*

Vowel	Recording	Condition	Mean	Mean difference	*t*	*df*	*p*
/u/	Words	Audio-visual 1	3.572	0.295	1.231	12	0.242
		Audio-visual 2	3.277				
/y/	Words	Audio-visual 1	3.661	-0.116	-0.287	12	0.779
		Audio-visual 2	3.777				
/u/	X-words	Audio-visual 1	3.629	-0.001	-0.002	12	0.998
		Audio-visual 2	3.629				
/y/	X-words	Audio-visual 1	3.700	0.472	0.877	12	0.397
		Audio-visual 2	3.229				

'X-words' refers to extracted words

Table 7. *Mean Ratings and Mean Difference in Mean Ratings at the Delayed Post-test Level for Isolated Words and Extracted Words in the Audio-visual Conditions in Experiment 1 (Audio-visual 1) and Experiment 2 (Audio-visual 2)*

Vowel	Recording	Condition	Mean	Mean difference	*t*	*df*	*p*
/u/	Words	Audio-visual 1	2.960	-0.602	-1.589	12	.138
		Audio-visual 2	3.562				
/y/	Words	Audio-visual 1	3.646	0.007	.012	12	.991
		Audio-visual 2	3.639				
/u/	X-words	Audio-visual 1	3.133	-0.419	-1.472	12	.167
		Audio-visual 2	3.552				
/y/	X-words	Audio-visual 1	3.391	-0.286	-.569	12	.580
		Audio-visual 2	3.677				

'X-words' refers to extracted words

DISCUSSION AND CONCLUSION

The results of the present study suggest that simplified visual feedback generated from spectrograms does not significantly contribute to pronunciation improvement of /u/ and /y/, regardless of whether they are presented and studied individually or side-by-side. The comments from the feedback questionnaire that subjects completed at the end of the experiment may shed some light on these results.

A large majority of subjects welcomed the use of computer-generated visual feedback while studying pronunciation (the sessions were rated 5.86 on a scale from 1 to 7). They also highly appreciated the oral component absent in the original study: oral presentations of the sounds' articulatory characteristics, oral explanations for the visual displays, and subject-instructor repeated interactions. Indeed, they found it considerably more helpful than the accompanying written presentations and explanations, which they reported using only as reference. Subjects also highlighted the benefits of the visual displays which provided immediate feedback on their production[xiii] and heightened their awareness of articulator position,[xiv] allowing them to compensate for perception difficulties.[xv] The subjects in the second experiment also found it very helpful to study the visual representation of /u/ and /y/ side-by-side as it enabled them to pinpoint the exact problem on which to focus.[xvi]

While all the subjects receiving visual feedback soon found it comprehensible, a minority of them expressed difficulty understanding and interpreting the visual displays when

these were first introduced. Although the time allocated for training on the visual displays was double that in the original study,[xvii] it appears that this type of visual feedback, even when simplified, requires much longer initial presentation for subjects to feel comfortably familiar (O'Brien, 2006). It also seems that three minutes of practice on each target word may not be sufficient for subjects to be able to fully benefit from both the information provided by the visual displays and the interactive feedback with the instructor. Indeed, 9 of the 14 subjects who received visual feedback would have liked to have had more time to practice their pronunciation of each target word after interpretation of the visual display.[xviii] The importance of practice time has been stressed, with the recommendation to allow students to complete computer-based pronunciation practice beyond the classroom (Wang & Munro, 2004).

The benefits reported by the subjects seem to indicate that the experimental treatment in the present study partially meets the requirements of effective feedback: it does "not rely solely on the learner's own perception," it allows "verification of response correctness," and it pinpoints "specific errors" to the learner. The user's interpretation of the display, as well as interactive feedback from the instructor, provide opportunities to "suggest a remedy" (Neri et al., 2002). Additional time for familiarity with the visual displays, as well as for practice of sounds, may enhance the benefits of spectrographic visual feedback, which has often been considered hardly effective for pronunciation improvement at the segmental level (Neri et al., 2002). Moreover, several subjects pointed to the contribution of the visual displays to the development of perception skills. Considering that training on sound discrimination alone has had a positive impact on production (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997; Coniam, 2002; Levis, 2007), and that phonemic distinction between /u/ and /y/, both at the perception and the production level, is contingent on time and experience (Flege, 1987), the effect of these visual displays at the perception level in the long-term would also merit close examination before this type of feedback is definitively considered ineffective. Future research comparing the effect of visual feedback on the pronunciation of learners of different age categories, L2 proficiencies and/or learning styles is also warranted.

ABOUT THE AUTHOR

Viviane Ruellot began teaching French and applied linguistics at Western Michigan University in 2005. She serves as coordinator of beginning French courses and as liaison between WMU and Kalamazoo high schools. Her research focuses on the acquisition of pronunciation, more particularly that of French as a foreign/second language (L2). She looks for ways in which visual feedback may help learners bridge the gap between perception and production and improve their pronunciation. One of her goals is to find ways to simplify that information to make it available to learners not trained in speech analysis. She is also interested in the pedagogy of pronunciation teaching, as well as the stages of development in the acquisition of L2 French pronunciation.

[ix] Microsoft ® Sound Recorder, version 5.1.
[x] I.e., F1 (red): "the openness line", F2 (green): "the tongue position line", and F3 (blue): "the lip roundness line".
[xi] A short distance between the red (F1) and the green (F2) lines signifies that the tongue is pushed towards the back of the mouth, which is characteristic of /u/. Contrastingly, a wide distance between the same lines indicates the tongue is massed in the front of the mouth, which is necessary for the pronunciation of /y/.
[xii] Subjects were instructed not to pronounce the final consonant letter in *loup*, *nous*, *sous*, *fût*, and *nus*.
[xiii] "I feel that due to the immediate feedback received by the programs and being able to see what my pronunciation was, I was able to correct the sounds easily."
[xiv] "I was able to see [the] placement [of my articulators] and compare it to my own. I was able to see my mistakes and hone in on them." The visual displays presented side-by-side contributed to raising awareness of articulators and their position for a majority of subjects: "Before this class, I had never noticed a difference between those (…) sounds, and listening to and hearing oral explanations and examples helped, but WaveSurfer helped me to put my articulators where they needed to be."
[xv] "I have a hard time hearing the difference in my pronunciations so I used the 3-line model to see if I was doing it correctly;" "Especially with the /u-y/ practice I felt like I could really hear and feel the difference," and "I can now distinctly tell the difference between the sounds. Before I couldn't very well."
[xvi] "I was able to tell what exactly I needed to fix, like having my tongue closer to my teeth."
[xvii] In the original study, training on the visual displays amounted to 12 minutes. In the present study, it lasted for 24 minutes (i.e., 6 minutes of initial introduction in addition to the 18 minutes devoted to the study of the first target word).
[xviii] "These sessions were very helpful, but overall the class felt rushed because we crammed in the worksheets so that we could get enough pronunciation time."

REFERENCES

Bongaerts, T., Planken, B., & Schils, E. (1995). Can late learners attain a native accent in a foreign language? A test of the Critical Period Hypothesis. In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition: a critical look at the Critical Period Hypothesis* (pp. 30-50). Clevedon: Multilingual Matters.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101*, 2299–2310.

CALLIOPE (1989). *La parole et son traitement automatique*. J.P. Tubach, (ed.). Paris: Masson.

Coniam, D. (2002). Technology as an awareness raising tool for sensitising teachers to features of stress and rhythm in English. *Language Awareness, 11*(1), 30–42.

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect*, *18*(3), 53–70.

de Bot, K. (1983). Visual feedback of intonation I: Effectiveness and induced practice behavior. *Language and Speech, 26*(4), 331-350.

Derwing, T. M. & Munro, M. J. (2005). Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly 29* (3), 379-397.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*, 393-410.

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), 1-17.

Engwall, O., Wik P., Beskow, J., & Granström, B. (2004). Design strategies for a virtual language tutor. *ICSLP 2004* (3), 1693-1696. Retrieved August 24, 2010, from http://www.speech.kth.se/ville/publications/icslp2004_tutor.pdf

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.

Guilloteau, N. C. (1998). Modification of phonetic categories in French as a second language: Experimental studies with conventional and computer-based intervention methods. *Dissertation Abstracts International, A: The Humanities and Social Sciences, 58* (7, Jan, 2621-A).

Hardison, D. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology, 8*(1), 34-52.

Jenkins, J. (2004). Research in teaching pronunciation and intonation. *Annual Review of Applied Linguistics, 24*, 109-125.

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

Liljencrants, J., & Lindblom, B. (1972). Numerical simulations of vowel quality systems: the role of perceptual contrast. *Language, 48*, 839-862.

Lindblom, B. (1986a). Phonetic universals in vowel systems. In J. J. Ohala & J. J. Jaeger (Eds.), *Experimental Phonology* (pp. 13-44). New York: Academic Press.

Lindblom, B. (1986b). On the origin and purpose of discreteness and invariance in sound patterns. In J. Perkell & D. Klatt (Eds.), *Invariance and variability of speech processes* (pp. 493-510). New Jersey: Lawrence Erlbaum.

Lindblom, B. (1990). On the notion of possible speech sound. *Journal of Phonetics, 18*, 135-152.

Lindblom, B., & Maddieson, I. (1988). Phonetic universals in consonant systems. In L. M. Hyman & C. N. Li (Eds.), *Language, speech and mind* (pp. 62-78). London: Routledge.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Neri, A., Cucchiarini, C., & Strik, H. (2002). Feedback in computer assisted pronunciation training: Technology push or demand pull? *ICSLP-2002,* Denver, USA, pp. 1209-1212.

O'Brien, M. G. (2006). Teaching pronunciation and intonation with computer technology. In L. Ducate & N. Arnold (Eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching* (CALICO Monograph Series, Vol. 5, pp. 127–148). San Marcos, TX: CALICO.

Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston, Inc.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255-287.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic clues. *The Modern Language Journal, 84* (3), 372–389.

Ruellot, V. (2007). *French pronunciation learning and computer-mediated visual feedback*. Unpublished PhD dissertation, University of Illiois at Urbana-Champaign. DAI-A 67/07, p. 2556, Jan 2007, 3223704.

Sjölander, K., & Beskow, J. (2000-2005). WaveSurfer (Version 1.8.5) [Computer software]. KTH, Stockholm. Available: http://www.speech.kth.se/wavesurfer/download.html

Wang, X. & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*, 539–552.