PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING

Conference Dates and Location

September 19-21, 2013
Iowa State University
Ames, Iowa, USA

5th Annual Proceedings

EDITORS

John Levis, Iowa State University

Shannon McCrocklin, University of Texas Pan American

PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING
5th Annual Proceedings

Table of Contents

New approaches

Teaching Tips

Levis, J., & McCrocklin, S. (2014). Turning a Corner. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1-10). Ames, IA: Iowa State University.

Turning a Corner

John Levis, Iowa State University
Shannon McCrocklin, University of Texas, Pan American

After a year away, the 5th Pronunciation in Second Language Learning and Teaching Conference returned to Iowa State University on September 19-21, 2013. The theme was Pronunciation in the Language Teaching Curriculum. The conference drew 125 participants from 18 countries and 15 US states. From the perspective of many at the conference, it was the best yet in terms of quality and variety of the sessions, and in terms of opportunities to network. The conference seems to have turned a corner. It is a conference that many L2 pronunciation researchers now see as essential for learning about the latest research and for connecting to other researchers in the field. The conference has grown to include a focus on a wide variety of languages with sessions on German, Chinese, Spanish, French, Japanese and Swedish, as well as English at last year's conference. Although English-focused sessions still dominate overall numbers (perhaps not surprising for a North America based conference), the greater number of languages considered in the conference is critical for the field, which needs to consider important questions about L2 pronunciation from the perspective of many languages.

Lynda Yates (chair of the Department of Linguistics, Macquarie University, Australia) gave the plenary address on Friday morning. The title was **Learning how to speak: Pronunciation, pragmatics and practicalities in the classroom and beyond.** The abstract for her talk is included below. For the second year in a row, the plenary talk was not written in full for the conference proceedings. This is actually a good thing, since Lynda's plenary is instead being published by *Language Teaching* (Cambridge University Press), one of the top journals in the field. The confidence of *Language Teaching's* editor, Graeme Porte, in the growing influence and quality of PSLLT is evident by his asking our plenary speakers to submit their papers two years in a row. We would loved to have included her talk in our own proceedings, but their talks (and our conference) will have a much higher profile in the pages of *Language Teaching*.

Plenary Abstract

It is beyond dispute that learners who want to develop good speaking skills in a language also need to develop good pronunciation, and yet research continues to report that pronunciation still has low visibility in the curriculum and is often treated as something of a poor relation in the classroom. Many teachers are still wary of pronunciation as a specialist area that is somehow separate from the other skills necessary for successful communication - an isolationist tendency that can make its consequent neglect in the curriculum and in teacher training programs only too easy.

In this plenary I go back to basics and focus on what it is that learners need to do outside the classroom with the language they are learning. Drawing on studies that have explored the lives and communicative needs of immigrants and international students, I will illustrate not only the importance of pronunciation in their lives, but also its close interrelationship with other spoken skills. I will then consider the implications for how we approach the teaching of pronunciation proactively as part of developing students' repertoire of speaking skills in the classroom and beyond.

The conference also included a pre-conference workshop, **Models, metaphors, and the evidence of spontaneous speech: A new relationship for pronunciation and listening.** Presented by Richard Cauldwell of speechinaction, the workshop presented a new approach to listening based on the reality and messiness of normal connected speech. Approximately 40 people attended the full-day workshop. The workshop description is reproduced below.

Pre-Conference Workshop

This workshop has the goal of improving the teaching of listening, by identifying and exploiting a new relationship between pronunciation activities and listening goals. New concepts and techniques (both high- and low-tech) will be illustrated. Participants will leave the workshop with new ideas to consider, and activities to use immediately in the classroom. The workshop will begin with thought-provoking theory, and end with the ruthlessly practical: but throughout there will be a constant reference to the evidence of recordings of spontaneous speech, and continual opportunities for suggestions and questions from participants.

Rationale

For pronunciation and speaking, we encourage clear intelligible speech. We present learners with a model of speech which is built around dictionary

pronunciations (citation forms) and rules of connected speech. We can think of the citation forms as greenhouse plants – they are isolated forms preceded and followed by a pause, with their component parts – vowels, consonants, syllables and stresses, all clearly present. The rules of connected speech – linking, elision, sentence stress, etc - can be thought of as guidelines for transplanting and arranging greenhouse plants into orderly pleasing arrangements in a garden. However, the greenhouse forms and the gardening guidelines are not appropriate for teaching listening. This is because the speech that learners encounter outside the classroom is more like jungle vegetation than garden or greenhouse plants, much wilder than the forms they encounter in the classroom. Such speech contains phenomena which are rarely seen in textbooks and words, like vegetation in the jungle, are blended into their neighbours in ways which are not predicted by the rules of connected speech. They are squeezed into bursts of the stream of speech, and it becomes difficult to recognise where one word begins and another ends, or indeed whether word-endings, syllables, or whole words have occurred at all. In class, we need to prepare students for their encounters with jungle listening, while continuing to promote intelligible pronunciation. This workshop will describe and explore ways of working on these separate but related goals.

Workshop Timetable

Part 1: Models and metaphors -The goals of listening and pronunciation are different. We need different models of speech for each goal. We have good models in place for pronunciation, we have inadequate models for teaching listening. We need to distinguish between goals and pronunciation activities can serve the goal of listening.

Part 2: Evidence from spontaneous speech -Words have many different soundshapes, of which the citation form is only one. The soundshapes are formed by interactions between the language and speaker factors: gender, accent, choices of speed, prominence and clarity.

Part 3: High-tech solutions: computers, smartphones, tablets, etc. - Recent developments in technology enable us to examine what happens to words in the stream of speech, to compare how words sound different as speakers and contexts change. We can manipulate and play with the sound substance of speech, in ways which promote faster learning of the listening skill.

Part 4: Low-tech solutions: teachers and learners voices in the classroom -The teacher voice and student voices can together be used in class to create, savour and handle the sound substance of the stream of speech. We will look at a number of activities that can be used and adapted to different teaching contexts.

Sessions

In addition to the plenary and preconference workshop, there were 31 presentations at the conference, with two concurrent sessions Friday and Saturday. In addition, there were 22 posters, 9 Teaching Tips, and a conference dinner open to all participants. As the Teaching Tips presentations were put into action for this first time this year and are less common at scholarly conferences, they need further explanation. Part of the goal of PSLLT is to connect practice, research and theory, and this means that both theory and practice need to be represented in the conference. PSLLT is primarily a research-oriented conference. It has a wide appeal to researchers from many areas and working in many languages. This does not mean, however, that the conference does not appeal to language teachers. Indeed, most participants are language teachers in part of their professional lives. As a result, we started the Teaching Tips Roundtable based on an idea John got from a Speech colleague. At the annual national conference for Speech Communication professionals, she told him that there is a section titled "My Great Idea" for teaching the basic course in speech. Since this started, it has become one of the best-attended sessions of the conference and provides an opportunity for presenters who are more teaching-oriented to show what they do and to connect theory and practice. We stole the idea and tried it at PSLLT. Each presenter sat at a round table with 9 chairs. For 10 minutes, they demonstrated their teaching tip to a full table, giving out a handout and taking questions with any extra time. Then a bell rang and everyone was free to go to another table. Each presenter then had a new table of participants for another 10 minutes. Teaching Tip presenters did this 7 times during the Roundtable session, and participants were able to go to 7 different teaching tips. The session got some of the highest ratings of any during the conference. The schedule and titles of presentations, posters and teaching tips is given below.

Friday, September 20th

8:00-8:50am	**Registration (Cardinal Room)**	
9:00-9:10	**Welcome (Cardinal Room)**	
9:10-10:10	**Plenary Address by Lynda Yates (Cardinal Room)**	
10:10-10:30	**Break**	
	Cardinal Room	**Gold Room**
10:30-10:55	**Erin Zimmerman** Teaching the Teachers: How Do Pronunciation Textbooks Aid Inexperienced Teachers' Pedagogy?	**Murray Munro** What do you know when you "know" an L2 vowel?
11:00-11:25	**Sinem Sonsaat & Stephanie Link** How do nonnative teachers use pronunciation materials? Implications for materials development	**Ron Thomson** Does vowel learning in one context generalize to other contexts?
11:30-11:55	**Ashley Rocammo** Learning Pronunciation in Just Ten Minutes a Day: Adapting	**Ettien Koffi** Assessment of the Intelligibility of [ʌ] in Seven Varieties of L2

	Pronunciation Training to a Four-Skills German Classroom	Englishes
12:00-1:50	**Working Box Lunch (Provided in Pioneer Room)**	
12:30-1:50	**Posters: Pioneer Room**	
	S. Alexander -Intonation and perceived sincerity in EFL and ESL learner apologies **J. Barcroft & M. Sommers –** Better L2 pronunciation is one of the many benefits of acoustically varied input **C. Barrett –** Laying a foundation for rhythm-based pronunciation instruction **C. Cárdenas**- Scaphoning your language **S. Chibani**- Pronunciation teaching in Algeria: From stagnation to progress **L. Cai –** An efficient method to build up native sounds in Chinese teaching: Multi-sensory and multi-cognitive approaches **M. Delicado Cantero & W. Steed –** Fair Dinkum: L2 Spanish in Australia by the book **F. Desmeules-Trudel-** VISC effects on the perception of Quebec French nasal vowels by Brazilian learners **N. Driscoll –** Hatsuon Help: a research-based, culturally-sensitive English pronunciation website for Japanese ELLs **V. Gonzalez Lopez & D. Counselman**- The production and perception of Spanish voiceless stops by novice learners: shedding light on early L2 category formation **S. Halicki –** Back door phonetic conditioning: Accent therapy in early French pronunciation training **Y. Lan –** Detecting L2 speech deviations by a communicative experiment procedure: taking Cantonese speakers' realizations of English [r] as an example **S. Link, S. Sonsaat, & J. Levis –** Confidence in teaching pronunciation: How native and nonnative teachers negotiate the pronunciation classroom **W. McCartan –** Word stress diagnostic procedure shared through a wiki site **C. Nagle -** Acquisition of the voicing contrast in L2 Spanish **D. Olson & H. Offerman**- The effects of visual feedback on learner pronunciation: Speech analysis software in the L2 classroom **L. Pierce –** Multi-methodological, cross-disciplinary approaches to pronunciation teaching **S. Shoji –** Japanese epenthetic vowels: How Japanese speakers pronounce English words **K. Taylor de Caballero & S. Thompson-** Coloring pronunciation across the curriculum with the Color Vowel Chart	

	H. Yang – Investigating needs of stakeholders of an oral proficiency test for ITAs to bridge the gaps between ITAs' needs and raters' feedback. **E. Zetterholm –** Final stops or not? The importance of final consonants for an intelligible accent. **E. Zetterholm & M. Tronnier –** Different stress patterns meet: Kurdish L1 speakers learn Swedish	
	Cardinal Room	**Gold Room**
2:00- 2:25	**Larissa Buss** Beliefs and Practices of Brazilian EFL Teachers Regarding Pronunciation	**Takehiko Makino** Pronunciation Characteristics of Japanese Speakers' English: A Preliminary Corpus-Based Study
2:30-2:55	**Veronica Sardegna** Non-Native Teachers' Identity Formation as Qualified Pronunciation Teachers	**John Esling** The two-part model of the vocal tract: a new articulatory basis for phonetics
3:00-3:25	**John Levis, Stephanie Link, Sinem Sonsaat, Taylor Anne Barriuso** Native and nonnative teachers of pronunciation: Does language background make a difference in learner performance?	**Jessica Sturm** Effects of Instruction on Voice Onset Time in word-initial /p/ for L1 American English students: A Preliminary Study
3:30-3:55	**Break**	
	Cardinal Room	**Gold Room**
4:00-4:25	**Shannon McCrocklin –** Dictation Programs for Pronunciation Learner Empowerment	**Jacques Koreman, Olaf Husby, E. Albertsen, P. Wik, A. Øvregaard, & S. Nefzaoui** L1 variation in foreign language teaching: challenges and solutions
4:30-4:55	**Joan Sereno, Larry Lammers, & Allard Jongman** Perception of foreign-accented speech	**Patricia Watts, Amanda Huensch, & Lisa Pierce** Attainable Targets for L2 Learners: How Proficient L2 Speakers can Bridge the Gap
6:00	**Conference Dinner at St. Johns (See Map & Directions on Page 37)**	

Saturday, September 21st

8:30-9:00	**Registration (Cardinal Room)**
9:00-10:30	**Teaching Tips Round Robin (Cardinal Room)**

	C. Keppie - From Mirrors to Mouthwash: Instructional Approaches to Teaching Pronunciation **G.M. Levis & J. Levis** – Using introductions to improve initial intelligibility **C. Meyers**- Intelligible Accented Speakers as Pronunciation Models **A. Saalfeld** - Flipping the phonetics classroom **M. Reed** - The English syllable: Big news, bad news, and why it's important for intelligibility **M. Richards**- Providing individualized homework and accountability for ITAs via Internet and LMS resources **A. Roccamo** - Effective Pronunciation Instruction in Beginner and Intermediate Language Classrooms **V. Ruellot** - Introducing French Nasal Vowels at the Beginner Level **S. Zhang** - Using Tongue Twisters to Supplement CFL students' Pronunciation and Tone Practice	
10:30-10:55	**Break**	
	Cardinal Room	**Gold Room**
11:00-11:25	**Anne Violin-Wigent** Comparing online vs. face-to-face classes: A case study of a French pronunciation class	**Talia Isaacs, Jennifer Foote, & Pavel Trofimovich** Drawing on teachers' perceptions to adapt and refine a pedagogically-oriented comprehensibility scale for use on university campuses
11:30-11:55	**Jennifer Foote & G. Smith** Is there an App for that? An investigation of pronunciation teaching apps	**Murray Munro, Tracey Derwing, Ron Thomson, & Diane Elliot** Naturalistic L2 Segment Development: Implications for Pedagogy
12:00-12:30	**Moonyoung Park & Sarah Huffman** The Potential of ASR for Non-native English Speakers in Air Traffic Control	**Beth Zielinski** Demystifying comprehensibility for the language teaching curriculum
12:30-2:00	**Lunch (not provided)**	
	Cardinal Room	**Gold Room**
2:00- 2:25	**Yuan Zhuang** Suprasegmentals and second language teaching: A meta-analysis	**Marnie Reed** Connecting pronunciation to listening: Raising learner and instructor awareness
2:30-2:55	~~**Okim Kang & F. Chowdhury**~~ ~~Prosodic Features in L2~~	**Mari Sakai & Christine Moorman** Can perceptual training improve

	~~Accented Speech: Human versus Machine~~	production of L2 phones? A meta-analytic review
3:00-3:25	**Shelley Staples** Prosodic patterns in nurse-patient interactions: a comparison of international and U.S. nurses	**Richard Cauldwell** Pronunciation and Listening, the need for two separate models of speech
3:30-3:55	**Break**	
	Cardinal Room	**Gold Room**
4:00-4:25	**Ghinwa Alameen** Perception and production of Linking in Non-Native Speakers of English	**Shannon McCrocklin & Stephanie Link** What is identity? ESL and Bilinguals' Views on the Role of Accent
4:30-4:55	**Paul Keyworth** The Acoustic Correlates of Stress-shifting Suffixes in Native and Nonnative English	**Christina Shea & Jennifer Vojtko Rubi** Dialect adaptation and L2 Spanish listeners
5:00-5:30	**Closing (Cardinal Room)**	

The 5th Annual Proceedings

The PSLLT Proceedings this year features 20 papers. Because L2 pronunciation is becoming a hotter topic and many journals are publishing papers that feature L2 pronunciation, a large number of presenters told us that they were planning to submit their papers to peer-refereed journals, including the new *Journal of Second Language Pronunciation* (John Benjamins), a journal that is a direct outgrowth of the PSLLT conference and its electronic proceedings. This is good news. Established peer-reviewed journals have high impact on the field, and the goal of the proceedings has always been to increase the impact of L2 pronunciation research. This seems to be happening, and we are happy that the proceedings are being joined by a larger number of publications on L2 pronunciation.

Proceedings papers come from the presentations, posters and Teaching Tips. Seven of the proceedings papers are from the Teaching Tips section. There is no clear venue for such papers, and we are especially pleased to make the write-ups from these very popular sessions available to the field.

PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING
5th Annual Proceedings

Table of Contents

New approaches

Teaching Tips

Future Conferences

The next three conferences are also scheduled. They will be

6th annual conference - September 2014 University of California at Santa Barbara

7th annual conference – October 2015 at Texas A&M, Dallas, Texas

8th annual conference – August or September, 2016, Calgary, Alberta, Canada

For information, go the conference website at www.psllt.org

Koffi, E. (2014). An instrumental account of the intelligibility of [ʌ] in seven varieties of L2 Englishes. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 11-21). Ames, IA: Iowa State University.

An Instrumental Account of the Intelligibility of [ʌ] in Seven Varieties of L2 Englishes

Ettien Koffi, St. Cloud State University

Munro, Flege, and MacKay (1996, p. 328) and Munro and Derwing (2008, p. 493) report the results of perception studies in which they found that [ʌ] was one of the least well perceived vowels by General American English (GAE) hearers of L2 Englishes. Exploratory acoustic phonetic studies conducted on seven varieties of L2 Englishes support their findings in part. Indeed, the vowel [ʌ] in these seven varieties of L2 Englishes overlaps acoustically with or encroaches on [æ] or [ɑ]. As a result, GAE hearers may have a hard time perceiving [ʌ] accurately. However, confusion data from Peterson and Barney (1952) and Hillenbrand et al (1995) also indicate that [ʌ] is among the least well perceived vowels of GAE. It is perceived accurately 92.2% of the time in Peterson and Barney, and 90.8% of the time in Hillenbrand et al. The infelicitous perceptions of [ʌ] may be due to the realignment of vowels in the acoustic vowel space that is going on presently in GAE. As a result, some other vowels are overlapping with the acoustic vowel space of [ʌ]. Small (2005, p. 79) notes, for instance, that many participants in his acoustic phonetic studies confuse [ʊ] and [ʌ]. I contend in this paper that the poor intelligibility of [ʌ] may have as much to do with the dialect(s) of the intelligibility judges as with the acoustic production of the L2 talkers. Furthermore, I contend that researchers can gain greater insights into the intelligibility of vowels if L2 production data is assessed instrumentally and used in tandem with confusion data that is already available for GAE and other accented Englishes. Doing so can help us determine the real sources of the intelligibility problems with L2-accented production of [ʌ].

Classificatory and Perceptual Difficulties

It is practically impossible to classify the vowel [ʌ] by itself without having to make a reference to another vowel. Therein lie the production and perception difficulties that will be addressed in this paper. Fromkin, Rodman and Hyams (2014, p. 206) define [ʌ] as follows: "The vowel [ʌ] in the word *luck* [lʌk] is a central vowel pronounced with the tongue low in the mouth though not as low as with [ɑ]." In this case, [ʌ] is contrasted with [ɑ]. Ladefoged (2006, pp. 90, 219) also defines [ʌ] by contrasting it with [ɔ] in one case, and [ɜ] in another case. In the "official" IPA chart, [ʌ] is classified as a back vowel where it occupies the same position with [ɔ]. These classificatory difficulties are symptomatic of the perception hurdles that GAE hearers face when they are asked to render intelligibility judgments on the segment [ʌ] produced by non-native speakers.

More often than not, they mistake non-native [ʌ] for [ɑ/ɔ] or for [æ]. In this paper, acoustic measurements of [ʌ] are provided in three dialects of American English and seven varieties of L2 Englishes to explain why these confusions exist.

Data Collection and Background Information

The data for the acoustic measurements of [ʌ] in L2 Englishes come from Arabic, Mandarin, Hispanic, Japanese, Korean, Slavic, and Somali speakers of English who were enrolled in my advanced undergraduate phonetics and my graduate phonology courses. The data from the three dialects of American English are from Peterson and Barney's (1952) classic study of GAE vowels, Hillenbrand et al.'s (1995) replication of their study for Midwestern English, and Koffi's (2013) replication of these two studies for the study of vowels in Central Minnesota. The non-native speaking participants in this study were asked to produce the same eleven words that native speakers of American English produced in the three studies mentioned above. The words are: *hid, heed, hayed, head, had, who'd, hood, hoed, hawed, hod,* and *hud*. Each word was produced three times. The words were recorded on laptop computers with built in microphones. Approval was obtained from the Institutional Review Board prior to the beginning of the study. The number of participants varies greatly, from three in the case of Slavic speakers to more than twenty in the case of Central Minnesota speakers.

The vowel [ʌ] is worth singling out for study for three main reasons. First, it is a high frequency vowel in English. According to Faircloth and Faircloth (1973, p. 18), it is the eighth most frequent vowel in English. It also carries a moderate relative functional load. According to Catford (1987, p. 89), the relative functional load of [ʌ] vs. [æ] is 68%, the one for [ʌ] vs. [ɑ/ɔ] is 65%, and [ʌ] vs. [ʊ] is 9%. The second reason for studying [ʌ] has to do with the fact that it is more prone to regional variations than any other GAE vowels. For this reason, I contend in this paper that some of the poor intelligibility scores given by intelligibility judges has as much to do with the judges' own inability to perceive [ʌ] accurately as with the inaccurate production by non-native speakers. Finally, [ʌ] is worth studying because its F1 formant values often overlap with those of [æ] or [ɑ/ɔ] in L2 Englishes.

The paper is organized as follows. Section 3.0 provides the necessary background for the instrumental assessment of intelligibility. In section 4.0, I provide acoustic data to support my contention that GAE talkers and hearers do not perceive [ʌ] completely accurately. It is customary in the acoustic phonetic study of vowels to discriminate between adult females and adult males because of the significant differences that are found in the laryngeal structures of the two genders. For this reason, the intelligibility of [ʌ] in L2 varieties of English is divided according to the gender of the participants. Section 5.0 focuses on the acoustic vowel space of the female participants, while section 6.0 concentrates of the acoustic vowel space of their male counterparts. For each gender group, cursory explanations are offered to assess the intelligibility of [ʌ] instrumentally. More in-depth discussions are devoted to how Mandarin females and Spanish-speaking males produce [ʌ]. Mandarin and Spanish-speaking [ʌ] are singled out for extra scrutiny

because there are already published studies on vowel confusion in these two varieties of L2 English. Focusing on them offers the opportunity to test the validity of an instrumental assessment of vowel intelligibility.

The Foundational Principles of Instrumental Assessment of Intelligibility

The overwhelming majority of the claims about the intelligibility of L2 vowels are based on how native speaker judges aurally perceive the vowels produced by non-native speakers. The present study departs from this long-held tradition and seeks to assess the intelligibility of L2 vowels instrumentally. In this approach, intelligibility is based purely on the acoustic cues produced by non-native speakers. At the core of this analysis is the Perceptual Distance Hypothesis (Johnson 2012, p. 119). It is formulated here as follows:

Perceptual Distance Hypothesis (PDH)
Segments that are acoustically closer tend to be confused with each other.

Following Labov et al. (2013, p. 43), I propose that if the acoustic distance between two front vowels or two back vowels is 60 Hz or less on the F1scale, intelligibility can be compromised. The calculation of the acoustic distance between vowels is based solely on the F1 formant frequency because, according to Ladefoged (2006, p. 188), it contains by itself 80% of the acoustic energy of the vowels. Generally though, since there are varying degrees of intelligibility (Byrd and Mintz 2010, p. 72), the smaller the acoustic distance between vowels, the greater the potential for confusion, i.e., unintelligibility.

The Vowel [ʌ] in Three Dialects of GAE

Before GAE judges can render reliable judgments on the intelligibility of [ʌ] in L2 English, we must make sure that they can perceive it accurately when it is produced by other GAE talkers. Available confusion data shows that [ʌ] is the second worst perceived vowel in English. According to Peterson and Barney (1952, p. 183, Table I), [ʌ] is confused with [ɑ/ɔ] 5% of the time. Overall, GAE hearers perceive it accurately 92.2% of the time. The percentage of accurate perception drops to 90.8% of the time in the Midwest (Hillenbrand et al. 1995, p. 3108, Table VI) because [ʌ] is confused with [ɑ/ɔ] 5.5% of the time. It is also confused with [ʊ] 3.2% of the time.1 In fact, Table 1 shows that there has been a gradual but steady reduction of the acoustic distance between [ʌ] and [ɑ/ɔ] on the one hand, and [ʌ] and [ʊ] on the other, in three dialects of American English.

1 The focus here is on female speech because I have not yet dealt with the acoustic vowel space of Central Minnesota male English. In Peterson and Barney, [ʌ] is confused with [æ] 0.07% of the time. In the Midwest, [ʌ] is not confused at all with [æ]. Other rates of confusion are so infinitesimal that they do not warrant any further comments.

Table 1
Comparative Formant Values in Three Dialects of Female US English

Words		hod	hud	hood	Distance	Distance
Vowels		[ɑ]	[ʌ]	[ʊ]	[ʌ] vs. [ɑ]	[ʌ] vs. [ʊ]
Peterson and Barney	F1	850	760	470	90	290
	F2	1220	1400	1160	180	240
Hillenbrand et al.	F1	936	753	519	183	234
	F2	1551	1426	1225	125	201
Central Minnesota	F1	855	743	626	112	117
	F2	1462	1643	1519	181	124

The acoustic distance between [ʌ] and [ʊ] deserves special attention because the two segments are being increasingly confused by American speakers. The acoustic distance between [ʌ] and [ʊ] is 290 Hz in Peterson and Barney in GAE female English (760 – 470 Hz). It drops to 234 Hz in Hillenbrand et al. (753-519 Hz). Among Central Minnesota female talkers, the acoustic distance drops precipitously to 117 Hz (743 – 626 Hz). Central Minnesota is far from being the only region of the USA where the acoustic distance between [ʌ] and [ʊ] has dropped so drastically. Small's (2005, p. 79) claim that "college students in the Midwest confuse [ʌ] and [ʊ]" is further evidence that this phenomenon may be far more widespread than has been acknowledged in the literature. Prator and Robinett (1985, pp. 138-140) also write about the confusion of [ʌ] and [ʊ]. I contend that the reduced acoustic distance between [ʌ] and [ʊ], as evidenced by the data on Central Minnesota English predisposes many speakers to not perceive [ʌ] and [ʊ] accurately. Data from Boberg (2008, pp. 137-139) and Walden (2012, pp. 188-189) on Canadian English indicate that speakers from the British Columbia area from where many of Munro and Derwing's (2008) intelligibility judges were recruited are predisposed at not perceiving the difference between [ʌ] (760 Hz) and [ʊ] (619 Hz) accurately because the acoustic distance between them in their dialect of Canadian English is only 141 Hz. Consequently, the intelligibility judgments that they render on L2 [ʌ] are more likely to be inaccurate. 2

The Intelligibility of [ʌ] in L2 Female Englishes

One main difference between L1 and L2 Englishes as far as confusion is concerned is that native speakers hardly confuse [ʌ] with [æ], while such confusions are commonplace in non-native varieties of English. Articulatorily, non-natives speakers pronounce [ʌ] in ways that are not confusable with [ʊ], unlike some native speakers of English. These two differences explain why [ʊ] is not mentioned in this section and the next. They also justify the inclusion of [æ] in these two sections, but not in the previous one. The native

2 During this research project, acquaintances of mine recounted many anecdotal stories in which they have misperceived words such as <hut> vs. <hot>, <mums> vs. <moms>, <buck> vs. <bock>, <cut> vs. <cot>, <gut> vs. <got>.

languages of the female talkers whose [ʌ] is studied here are Japanese, Korean, Mandarin, Slavic, and Spanish, as shown in Table 2.

Table 2
Comparative Formant Values in L2 Female Englishes

Words		had	hud	hod
Vowels		[æ]	[ʌ]	[ɑ]
Japanese	F1	844	894	772
	F2	1685	1574	1525
Korean	F1	717	634	817
	F2	1589	1136	1487
Mandarin	F1	900	938	853
	F2	1968	1664	1397
Slavic	F1	821	882	843
	F2	1937	1683	1443
Hispanic	F1	847	719	746
	F2	1773	1493	1446

The comparative acoustic vowel space in Figure 1 below is based on the information contained in Table 2.

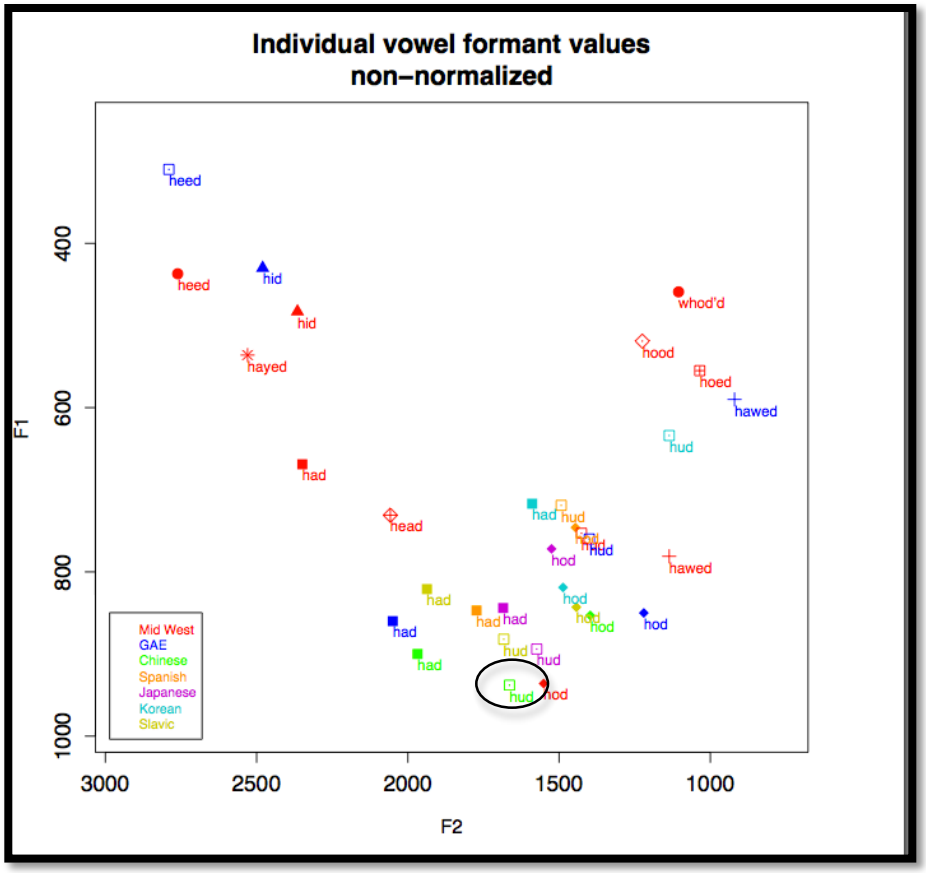


Figure 1. Comparative Acoustic Vowel Space of L2 Female Englishes

A look at Figure 1 shows that the Korean [ʌ] (634 Hz) can be very easily confused with [ɔ] (590 Hz) in GAE because the distance between them is 44 Hz. The Japanese [ʌ] (894 Hz) can be confused with [ɑ] (936 Hz) in GEA because they are separated by 42 Hz. The same is true for the Slavic [ʌ] (882 Hz) that is distant from [ɑ] in GEA by 39 Hz. The way in which Spanish-speakers produce [ʌ] (719 Hz) is barely distinguishable from the way they produce their [ɑ] (746 Hz) because only 27 Hz separate them. As is well known, human beings cannot perceive frequencies differences of less than 20 Hz (Ferrand 2007, p. 34). For this reason, it is doubtful that GAE hearers can unambiguously distinguish between [ʌ] and [ɑ] in Spanish-accented English (see section 6.0 for additional evidence). Since the acoustic distances between [ʌ] and [ɑ] or [ʌ] and [ɔ] in these L2 Englishes are less than 60 Hz, varying degrees of confusion (unintelligibility) are to be expected. We now leave these L2 varieties to concentrate on Mandarin-accented English because this L2 variety has benefited from three important vowel perception studies.

Jia, Strange, Wu, Collado, and Guan (2006) conducted a vowel intelligibility study involving 222 Mandarin speakers, and their findings provide empirical support for the Perceptual Distance Hypothesis discussed in 3.0. The researchers divided Mandarin talkers into three groups: 91 "Monolinguals," that is, people with little or no English, 77 "Recent Arrivals," namely people who had lived in the US less than two years, and 54

"Past Arrivals," those who had been in the USA between 3 and 5 years. They report the following intelligibility rating of their production of [ʌ]. This segment was perceived accurately at rates of 54.3% in the production of Monolinguals, 44.7% in the speech of by Recent Arrivals, and 48.8% in the samples provided by Past Arrivals. The results of this study suggest that length of residency (LOR) does not have any positive effect on the production of [ʌ]. My acoustic measurements of [ʌ] produced by seven TESL graduate students whose native language is Mandarin bear this out. The average F1 formant of [ʌ] in Mandarin-accented English is 938 Hz. According to Hillenbrand et al. (1995), this anglicized [ʌ] encroaches on the acoustic vowel space of [ɑ] (936 Hz) in the dialect of GAE spoken in the Midwest. The acoustic difference of 2 Hz between the two is sub-sonic (Ferrand 2007, p. 34), that is, under the threshold of human aural perception. In other words, when Mandarin talkers produce [ʌ], Midwestern hearers perceive it as [ɑ] instead. Unlike Munro and Derwing (2008, p. 488) who found that Mandarin speakers improved their [ʌ] from periods T1/T2 to periods T5/T6, Jia et al.'s (2006) findings suggest that LOR does not improve the production of [ʌ]. In fact, the Mandarin talkers who provided data for this study had had on average 12 years of formal instruction in English as a foreign language in China prior to coming to the US. By the time their speech samples were collected, they had been living and studying in the US for two years or more. The acoustic data also seem to indicate that formal linguistics training does not seem to have any appreciable positive impact on their production of [ʌ]. My instrumental acoustic findings are further bolstered by Lai's (2010, p. 171) confusion study of Mandarin-accented English vowels. It was found that [ʌ] was perceived accurately only 22% of the time, while 78% of the time, GAE hearers confused [ʌ] with [ɑ]. This is not at all surprising, since the acoustic distance between Mandarin-accented [ʌ] and Midwest [ɑ] is only 2 Hz.

The Intelligibility of [ʌ] in L2 Male English

Let's now turn our attention to the ways in which the male talkers in the study produced [ʌ].

Table 3
Comparative Formant Values in L2 Male Englishes

Words		had	hud	hod
Vowels		[æ]	[ʌ]	[ɑ]
Arabic	F1	665	611	685
	F2	1504	1243	1154
Japanese	F1	702	596	542
	F2	1694	1385	1150
Korean	F1	640	572	694
	F2	1224	1055	1291
Somali	F1	678	629	609
	F2	1674	1532	1339
Spanish	F1	647	531	593
	F2	1491	1529	1196

Table 4
Comparative Formant Values in Two Dialects of US Male English

Words		had	hud	hod
Vowels		[æ]	[ʌ]	[ɑ]
Peterson and Barney	F1	660	640	730
	F2	1720	1190	1090
Hillenbrand et al.	F1	588	623	768
	F2	1952	1200	1333

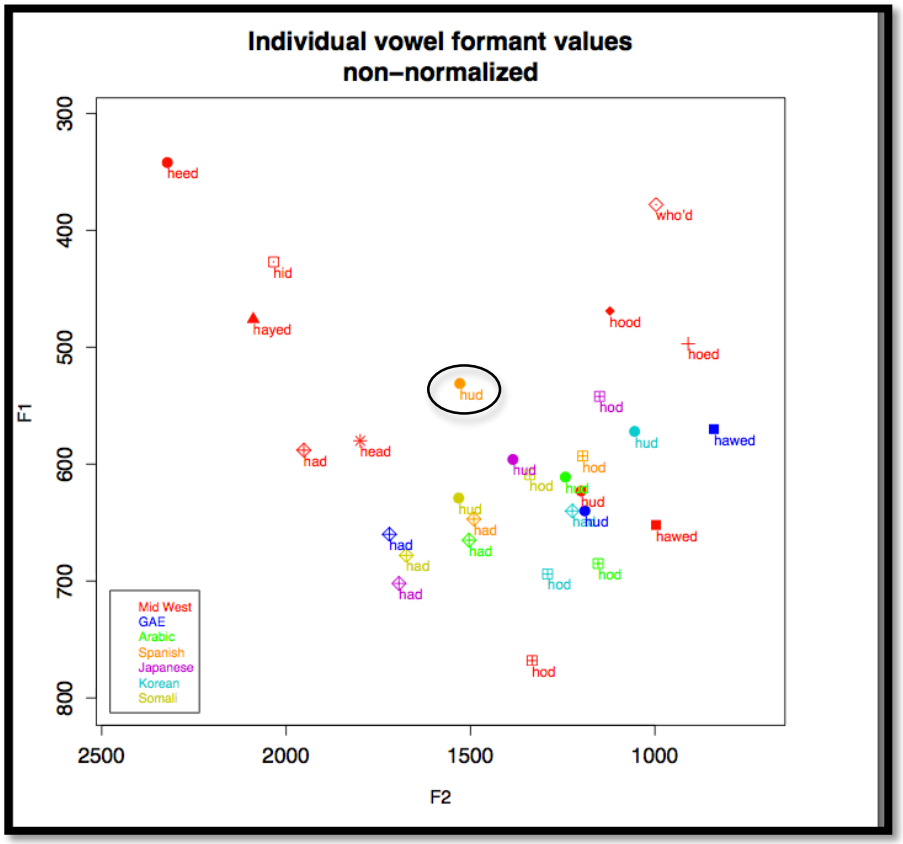


Figure 2. Comparative Acoustic Vowel Space for Males

Suffice it to say for now that Korean and Japanese-accented [ʌ] (572 Hz) may be misperceived by GAE hearers because it encroaches on the acoustic space of their [æ]. The segment [ʌ] produced by Arabic (611 Hz) and Somali (629 Hz) talkers is confusable with [æ] because their anglicized [ʌ] encroaches on the acoustic space of [æ] in GAE. Furthermore, the ways in which talkers produced [ʌ] in most of these varieties of L2 Englishes cause it to overlap acoustically with [æ], except in Japanese and Spanish. It can be concluded that intelligibility issues arise either because [ʌ] encroaches on [æ] in GAE, or because it overlaps with [æ] in accented Englishes, or both. Let's elaborate this further by examining the confusion data concerning [ʌ] in Spanish.

Lecumberri and Iragui (1997, p. 59) conducted a confusion study in which speakers of Received Pronunciation, a "standard" dialect of British English, were asked to rate the intelligibility of the [ʌ] produced by college students in Spain majoring in English. Overall, Spanish-accented [ʌ] was perceived accurately 81.6% of the time. However, the same study found that [ʌ] was confused with [æ] 15.8% of the time, and with [ɑ] 0.9% of the time. These findings lend support to the Perceptual Distance Hypothesis (PDH). Eight male speakers from various Latin American countries produced the [ʌ] circled in Figure 2. We see that their [ʌ] is the only one among the speakers of seven varieties of L2 English that is sufficiently distinct from other GAE vowels so as not to cause confusion. Yet, the acoustic data also shows the Hispanic [ʌ] (531 Hz) is acoustically not very distant from how Midwesterners produce [æ] (588 Hz). The acoustic distance of 57 Hz between these two segments means that confusion is still possible. The Spanish data confirm this by showing that [ʌ] is confused with [æ] 15.9% of the time. Not only does PDH predict accurately that confusion is likely between Hispanic-accented [ʌ] and [æ], but it also predicts accurately that GAE hearers are less likely to confuse Hispanic-accented [ʌ] with the way they produce [ɑ] (593 Hz). It is, therefore, not surprising that Lecumberri and Iragui found that [ʌ] was confused with [ɑ] only 0.9%.

Summary

This study complements the growing body of evidence that has been accumulating for more than a decade regarding the relatively poor intelligibility scores that GAE hearers give to L2 English pronunciations of [ʌ]. However, unlike the previous intelligibility studies that are based on impressionistic judgments rendered by native speaker judges, the present study has relied on an instrumental acoustic methodology to account for why non-native [ʌ] is often poorly perceived. In so doing, three contributing factors have been uncovered. The first has something to do with the dialect of American English that the intelligibility rater speaks. Data from Peterson and Barney (1952) and Hillenbrand et al. (1995) show that Midwest hearers perceive [ʌ] less accurately when they listen to GAE talkers from other regions of the USA. Acoustic phonetic data by Koffi (2013) find that Central Minnesotans fare even worse in their intelligibility of [ʌ] because, not only do they occasionally confuse it with [ɑ/ɔ] like other GAE speakers, but they are also increasingly confusing [ʌ] with [ʊ]. The second factor that contributes to the poor intelligibility of [ʌ] stems from how non-native speakers pronounce it. More often than not, they do not discriminate sufficiently in how they produce [ʌ], [æ], and [ɑ]. As a result, these segments overlap with each other in acoustic vowel space. Last but not least, L2 accented [ʌ] is often not intelligible because it encroaches on the acoustic vowel space of [æ] or [ɑ/ɔ] in GAE. Acoustic phonetic overlapping and/or encroachment interferes with intelligibility because in either case, GAE hearers cannot disentangle [ʌ] from [æ] or [ɑ] auditorily. Given the moderate to high relative functional load of [ʌ] and [æ] on the one hand, and [ʌ] and [ɑ] on the other, frequent instances of unintelligibility are expected to arise when non-native speakers produce monosyllabic lexical minimal pairs such <cup>, <cap>, and <cop> , or <but>, <bat> and <bought>.

ABOUT THE AUTHOR

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

REFERENCES

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

Boberg, C. (2008). Regional phonetic differentiation in standard Canadian English. *Journal of English Linguistics*, 36 (2), 129-154.

Byrd, D., & Toben, M. (2010). *Discovering speech, words, and mind*. Malden, MA: Wiley-Blackwell.

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

Fromkin, V., Rodman, R., & Hyams, N. (2014). *An introduction to language, 10th ed.*, New York: Cengage Learning.

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

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

Jia, G., Strange W., Wu Y., Collado, J., & Guan Q. 2006. Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America*, *119* (2), 1118-1130.

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

Koffi, E. (2013). The acoustic vowel space of central Minnesota English: Focus on female vowels. *Linguistic Portfolios*, 2: 2-16.

Labov, W., Rosenfelder, I. & Fruehwald, J. (2013). One hundred years of sound change in Philadelphia: Linear incrementation, reversal, and reanalysis. *Language*, 89 (1), 30-65.

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

Lai, Y. (2010). English Vowel Discrimination and Assimilation by Chinese-Speaking Learners of English. *Concentric: Studies in Linguistics*, 36 (2), 157-183.

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

Munro, M., & Derwing T. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning* 58 (3), 479-508.

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

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

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

Prator, C., & Robinett, B. (1985). *Manual of American English pronunciation, 4th edition.* New York: Harcourt Brace & Company.

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

Walden, M. (2012). Phonetic portfolio. *Linguistic Portfolios*, 1, 182-193.

Zetterholm, E. (2014). Final stops or not? The importance of final consonants for an intelligible accent. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 22-29). Ames, IA: Iowa State University.

FINAL STOPS OR NOT? THE IMPORTANCE OF FINAL CONSONANTS FOR AN INTELLIGIBLE ACCENT.

Elisabeth Zetterholm, Linnaeus University

To have a chance for integration in a new society as an immigrant you have to learn the language; it's a key to assimilation. A group of Karen speakers have lived in Sweden for some years but, unfortunately, it appears that they have great difficulties learning the pronunciation of Swedish in an intelligible way. There are phonetic and phonological differences between the Sgaw Karen language and Swedish. Sgaw Karen is described as having a three-way contrast for plosives in word-initial and word-final positions. There is also a monosyllabic structure and final stops or nasals are not pronounced. This causes problems when learning Swedish with a two-way contrast for initial plosives and heavy consonant clusters. A training program was constructed with sentences and minimal pairs of words ending with final consonants. Four learners imitated recordings of Swedish native speakers and recorded themselves in order to be aware of their own pronunciation. After six weeks of training they still had some problems with the final stops, but there was also some progression in their pronunciation, or at least an awareness of the problem.

INTRODUCTION

How well and how easy it is to acquire a new language often depends on many different things, both on an individual level as well as on a linguistic level, e.g., similarities and differences between the first and second language. For this study, second language learners of Swedish with a first language from Burma were recorded and analyzed. The two languages – Swedish and Sgaw Karen – are different in many ways, both concerning pronunciation and the alphabet. It appears that, still after eight years, these learners have great difficulties learning Swedish with an intelligible pronunciation and therefore have problems gaining access to the Swedish society. The Karen people in Sweden had been living in refugee camps in Thailand before arriving to Sweden. It is hard to tell if their experiences have any kind of impact on their ability to learn the new language.

Karen people in Sweden

Karen people are an ethnic group living in South East Asia, mostly in Burma (or Myanmar), southern China and the northern part of Thailand. Approximately 6 million Karen people live in Burma, but more than 100,000 live in refugee camps in Thailand. Since 2004, many Karen refugees have been resettled in Western countries, e.g., Sweden and USA, and families are spread all over the world. About 1,000 Karen people live in Sweden in small cities in the central and southern part of the country. Many of them, chiefly those middle-aged or older, have great difficulties learning Swedish at an intelligible level. That means that it is hard for them to get a job and access to the Swedish society. When arriving to Sweden they all have to study at SFI (Swedish for immigrants) but many of them do not manage to learn Swedish and after some semesters they often fail in their exams. Still after eight years in Sweden most of the adults cannot speak Swedish with an intelligible pronunciation. Therefore, it is a challenge to find out if there are specific problems related to their native language and to figure out how to teach them to speak Swedish. All of the Karen people, with few exceptions,

knowledge about, English. There has not been, until recently, any Swedish-Karen dictionary (http://karenswedishcommunity.org/Image/Information/Files/Swedish_Karendictionary.pdf). Those who have some knowledge about English often use both a Swedish-English dictionary and an English-Karen dictionary in order to try to learn Swedish. This is not an easy way to learn a new language. However, most of the young Karen speakers, ages 25 and below, have acquired the Swedish language and some are taking courses at Swedish universities. A few years ago the Karen people in Sweden established The Karen Swedish Community and one of the aims is to promote assimilation into Swedish society.

Karen languages

The Karen languages are a Tibeto-Burman branch from the Sino-Tibetan phylum consisting of many different dialects with influences from Burmese and other nearby languages. The number of Karen languages is not known, but there are probably 20-30 different languages, many of them not well documented. Languages spoken in the mountains usually have numerous dialects, some difficult to understand even for other Karen speakers (Manson, 2011). The languages are classified into three main groups; namely Northern, Central and Southern (Bradley, 1997). The two major dialects are called Sgaw and Pwo, both mainly spoken in the southern part of Burma with more than 1 million speakers of each of the two dialects. Sgaw Karen is a lingua franca among Karens (Naw, 2011). The letters of the alphabet derive from the Burmese characters with some modifications. The word order is SVO, which is uncommon among (mostly SOV) Tibeto-Burman languages (Naw, 2011).

The phonology of Sqaw Karen

The syllable structure in Karen languages is CV or C1(C2)V^{T}. C1 is any consonant, C2 is either a voiced velar fricative [ɣ], voiced bilabial approximant [w], voiced alveolar approximant [l] or a voiced alveolar trill [r]. V is a vowel and ^{T} is a tone. No words start with a vowel as a clear glottal stop always precedes a single vowel (Naw, 2011). Two varieties of the Sgaw dialect have been described by Jones (1961), the Moulmein Sgaw Karen and the Bassein Sgaw Karen. The phonological system is almost the same with 9 vowels and 27 or 23 consonants respectively (Jones, 1961), see Table 1 and Table 2. All consonants can occur in initial position, but syllable final consonants are not pronounced in Sgaw Karen due to historical developments (Manson, 2011). There are both aspirated and non-aspirated initial plosives [p^{h} p b, t^{h} t d k^{h} k], but no velar voiced plosive (Abramson, 1995; Jones, 1961). Moulmein Sgaw Karen has 6 different tones, two high, tow mid and two low, three of which have a glottal closure. Bassein Sgaw Karen has five tones. The tone system is based on voice quality, f0 and duration. Tone 1 is a slightly rising tone, while the other five tones are moderately falling. Voice quality, together with duration, seem to be the most important identification cues for listeners (Brunelle & Finkeldey, 2011).

Table 1
Sgaw Karen vowel inventory (Jones, 1961:62).

	Front	**Central**	**Back**
Close	i	ɨ	u
Mid	e	ə	o
Open	ɛ	a	ɔ

Table 2
Sgaw Karen consonant inventory, consonants only in Moulmein Karen in parentheses (Jones, 1961:62).

	Bilabial	**Dental**	**Alveolar**	**Postalveolar**	**Palatal**	**Velar**	**Glottal**
Plosive	pʰ		tʰ		(cʰ)	kʰ	
	p		t		(c)	k	ʔ
	b		d				
Fricative			sʰ				
		θ	s	ʃ		x	h/ɦ
			(z)			ɣ	
Nasal	m		n		ɲ	(ŋ)	
Lateral approximant			l				
			r				
Trill	w				j		
Approximant							

The phonology of Swedish
A brief description of Swedish phonology will highlight some areas of difference. For further information, see Bruce (2010) and Riad (2013). The syllable can consist of three initial and three final consonants at most, CCCVCCC. If so, the first consonant has to be an /s/ followed by a voiceless plosive and either of /l j r v/ in the beginning of a word. Any consonant can be in a final position. The vowel system consists of nine vowels (Table 3), each with a phonological distinction in quantity and variation in quality, which makes 18 distinct vowels in all. The duration is complementary which means that a syllable can only have one long segment, (V:C) or (VC:). There are 18 consonant phonemes, see Table 4. (Bruce, 2010).

Table 3
Swedish vowel inventory (Bruce, 2010).

	Front		**Back**
	Unrounded	**Rounded**	
Close	i	y	u
Mid	e	ʉ	o
Open	ɛ	ø	ɒ

Table 4
Swedish consonant inventory (Bruce, 2010).

	Bilabial	**Dental**	**Alveolar**	**Postalveolar**	**Palatal**	**Velar**	**Glottal**
Plosive	p		t			k	
	b		d			g	
Fricative		f	s		ç	ɧ	h
		v			j		
Nasal	m		n			ŋ	
Lateral approximant			l				
			r				
Trill							

A comparison of the two phonological systems of the languages show that there are differences concerning the phonemes which might be a problem for L2-learners of Swedish, e.g., the front rounded vowels and non-aspirated plosives; but the difference in the syllable structure indicate that epenthetic vowels or reductions will be used when learning to pronounce the Swedish consonant clusters and the last consonants. This is in accordance with earlier research about learning Swedish as a second language (e.g. Bannert, 2004).

Identified pronunciation problems for Sgaw Karen L1-speakers

The most obvious and general problem for Sgaw Karen speakers is the lack of pronunciation of the last consonant in Swedish words. For a native listener it is hard to hear word boundaries and it can change some words' meanings in Swedish. E.g., the word *tvål* [tvoːl] (soap) is pronounced *två* [tvoː] (two), *tåg* [toːg] (train) is pronounced *tå* [toː] (toe), *vit* [viːt] (white) is pronounced *vi* [viː] (we), *ljus* [jʉːs] (light) is pronounced *ju* [jʊ] (of course). Another perceptual problem is that the voiceless plosives without aspiration are often confused with a voiced plosive, /p-b, t-d/, e.g. *par* [pʰɑːr] (a couple) – *bar* [bɑːr] (bar), *tagg* [tʰag] (thorn) – *dagg* [dag] (dew). The meaning of the word changes if there is no aspiration on the initial voiceless plosive. All voiceless plosives are aspirated in stressed syllables in Swedish with one exception, namely if the plosive is preceded by an /s/.

THE STUDY

After identifying these specific pronunciation problems, a training session was constructed. The aim was to focus on final consonants and make the learners more conscious about the importance of the pronunciation in Swedish for a more intelligible accent. For this study, four speakers were selected from a database with recordings of Karen speakers at different places in Sweden. They had great difficulties and were willing to do study at home during the research. They received training sentences and two Swedish songs (details below) for exercising. They were recorded before and after the six weeks of training. There are also recordings, made by themselves, during the six weeks of training. The recordings were analyzed both with auditory and acoustic analyses using the program Praat.

The material for and instructions to the participants

Training sentences were constructed using words ending with a consonant some of which were linked together with the next word starting with a vowel or placed at the end of the sentence. The participants were told to practice by reading the sentences aloud, listening to two (one male and one female) Swedish native speakers and imitating and recording themselves using the program Praat (http//www.fon.hum.uva.nl/praat). They received basic instructions about how to do the recordings and listen to them. Some minimal pairs consisting of words ending with the plosives [g k] and different nasals [m n ŋ] were also constructed for practice, and were recorded by two native speakers for listening and imitation. In addition, they received two Swedish songs to learn. The two songs were chosen because they contained many common words ending with a consonant. The melodies were easy to learn and they had recordings to use for sing-a-long. If possible, they were requested to do the training together with each other, with Swedish friends or with help from their own children. Some of the Swedish friends got information and instructions about the study as well. The training sessions were not like a teaching situation, but more like self-study. They were also

All four participants were recorded when chatting to the research leader in order to get some spontaneous speech; when reading the sentences; and also when describing some pictures (a phoneme test) in the initial phase of the study and again after six weeks of training. To get an idea of the progression they were told to do recordings of their speech at least once a week. Some of them did that more often. They used Praat and a headset on their own computer at home.

To get an idea about the training sentences one example and a short explanation will be given: In the sentence: *Att åka tåg är kul* [atoːkaˈtoːgækʉːl] (going by train is fun) the aim is to link the last consonant in the first word /t/ and pronounce it like an initial consonant of the second word [toːka] and the final consonant in the third word will be pronounced as an initial consonant of the next word [gæːr]. The two 'new' words do not mean anything in Swedish but were meant to remind the speaker of the importance of the speech sound in Swedish.

The distinction between final consonants is of importance in some minimal pairs in the test and here are a few examples where tonal and non-tonal plosives as well as different nasals are in focus: *trygg* [tryg] (safe) – *tryck* [tryk] (press), *kam* [kam] (crest) – *kan* [kan] (be able to), *tunn* [tʉn] (thin) – *tung* [tʉŋ] (heavy).

Participants

Two male and two female Karen speakers, 45-55 years old, participated in this study. They were all born in Myanmar, had lived in refugee camps in Thailand for many years and in Sweden for eight years. They spoke the Sgaw Karen language, Burmese and some English. Even though they had lived in Thailand for many years they did not speak Thai. Their educational backgrounds differed a lot. One of the female speakers had been in school for 12 years, one of the male speakers only for three years as a child in Burma. All of them could read and write in the Karen language.

Analyses

The auditory analyses confirm that the pattern concerning the lack of pronunciation of the last consonants is a general problem for all speakers, e.g., *gaffel* [gafəl] (fork) is pronounced [gafə]. Figure 1 is an example of a female speaker's first try, to the left, and after instructions from a Swedish speaker and imitation, to the right. This is a clear pattern especially in the spontaneous speech and the phoneme test. When reading the sentences all informants sometimes pronounced the consonants and perhaps the spelling was a clue to remember to do so. In the spontaneous speech they often add an epenthetic final vowel after the consonant in order to end with a vowel and get close to the CV-structure of their first language. E.g. *stol* [stuːl] (chair) is pronounced [stuːlə], *näbb* [nɛb] (beak) is pronounced [nɛbɛ], bok [buːk] (book) is pronounced [buːkə].

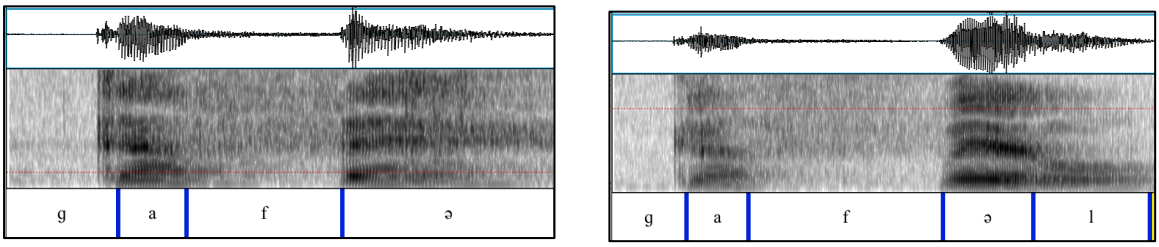


Figure 1. The word *gaffel* [gafəl] (fork) is pronounced [gafə] to the left and correctly to the right.

Voiceless aspirated plosives between voiced segments, e.g., vowels, are often pronounced without aspiration and sound like a voiced plosive for a native listener, e.g. *apa* [ɒːpʰa] (ape) is pronounced like [ɒːba]. Another observation in the training session and in some of the recordings is when one of the male speakers exaggerated the final plosives to make a distinction between a voiced and a voiceless final consonant in a minimal pair, *vig* [viːg] (lithe) and *vik* [viːkʰ] (bay). But, as seen in Figure 2, he exaggerated the pronunciation and made the consonant voiceless in both words. The difference is chiefly in the duration, especially the vowel duration.

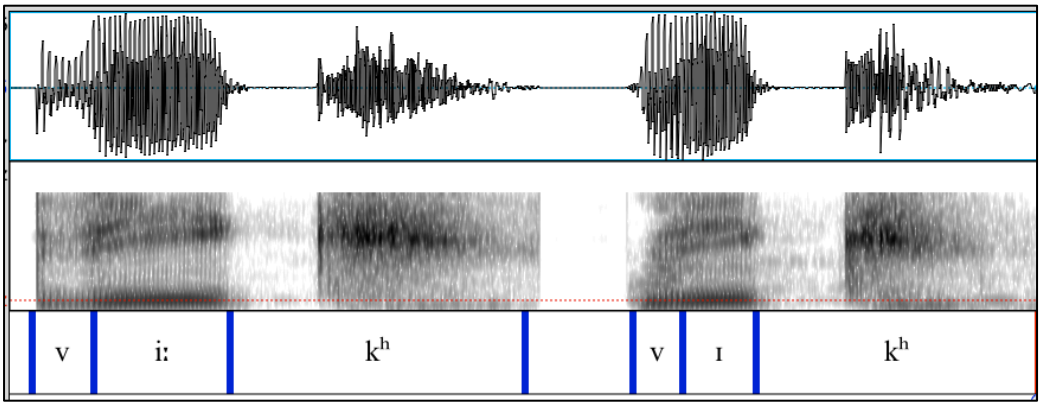


Figure 2. The minimal pairs *vig* [viːg] (lithe) to the left and *vik* [viːkʰ] (bay) to the right.

However, there were some differences in the recordings before, during and after the training for all speakers though in general they still had problems with the final consonants. One female speaker did not pronounce the final consonants in the words like *mjölk* [mjølk] (milk), *nål* [noːl] (needle), *rädd* [rɛd] (afraid) even after several trials in the recording after training. The other female speaker was quite successful and the two figures, 3 and 4, are examples of her progression in the pronunciation after six weeks of training. She is eager to learn Swedish and can discuss the language at a metalinguistic level. Unfortunately, all of the participants were not that successful.

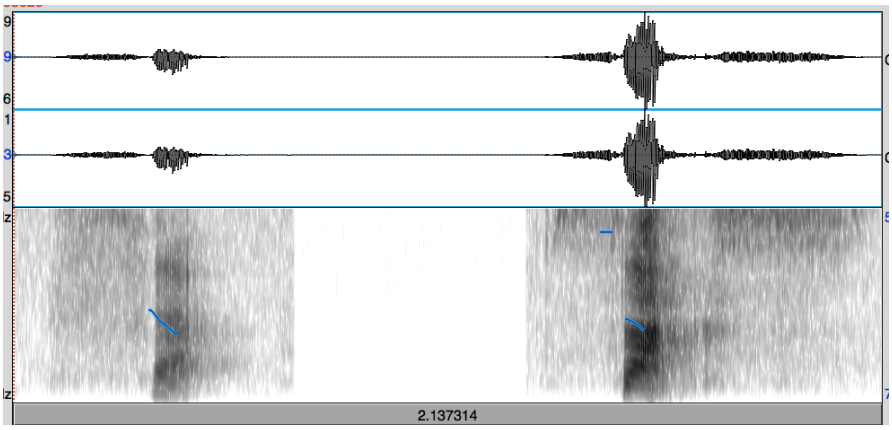


Figure 3. The word *sax* [saks] (scissor) before (to the left) and after (to the right) training.

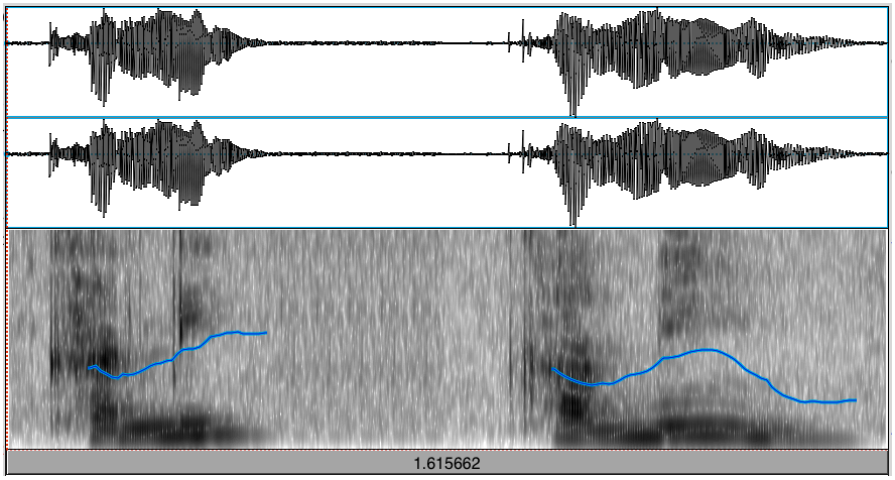


Figure 4. The word *kanin* [kani:n] (rabbit) before (to the left) and after (to the right) training.

CONCLUDING REMARKS

After six weeks of training the speakers still had some problems with their pronunciation in Swedish. They reported that their Swedish friends had been very helpful and had pushed them to do their training sessions and recordings. The participants' training schedules and recordings confirmed that they had been diligent and serious in their aim to learn and imitate the Swedish speakers.

Even though the participants still did not pronounce the final consonants in every word, it seemed that they were more aware of the problem. They were eager to learn Swedish, seemed to have a metalinguistic awareness and were able to discuss this phenomenon. To hear the sounds and correct pronunciation when listening to native speakers and to be aware of specific difficulties when learning a new language was a good starting point. The next step must be training, individual as well as together with someone who is able to give feedback. To do recordings of oneself, and listen to it, was another useful way to learn the pronunciation of a language. The results of this study show that the idea of training with linking words and minimal pairs can be successful for speakers who have an open syllable structure, CV, in their first language and have to learn a language with another more complicated syllable structure with final consonant clusters.

ACKNOWLEDGEMENTS

The research was partly supported by Birgit and Gad Rausings Foundation, Åke Wibergs Foundation and Magnus Bergvalls Foundation, Sweden. Special thanks to Eva and Lennart Boström who introduced me to the Karen People and to all the Karen speakers I've met so far.

ABOUT THE AUTHOR

Elisabeth Zetterholm is a senior lecturer and researcher in Swedish as a second language at Linnaeus University, Sweden. Her current research interests focus on pronunciation in second language acquisition and how to learn to speak Swedish in a way to be understood. When teaching, most of her students are prospective teachers of Swedish as a second language and therefore she wants to develop the teaching methodology. She received her PhD in phonetics at Lund University, Sweden (2003). Her thesis and postdoctoral research were about voice imitation, with a focus on speaker identification and individual features in voice and speech.

Elisabeth Zetterholm, PhD
Department of Swedish
Linnaeus University
SE-351 95 Växjö
Sweden
Tel: +46 70 5354183
Email: elisabeth.zetterholm@lnu.se

REFERENCES

Abramson, A.S. (1995). Laryngeal timing in Karen obstruents. In *Producing speech: Contemporary Issues*. For Katherine Safford Harris. F. Bell-Berti, & L.J. Raphael (eds). (pp. 155-165). AIP Press: New York.

Bannert, R. (2004). *På väg mot svenskt uttal*. Studentlitteratur.

Bradley, D. (1997). Tibeto-Burman languages and classification. In *Papers in Southeast Asian Linguistics 14*. *Pacific Linguistics A, 86*, 1-71.

Bruce, G. (2010). *Vår fonetiska geografi. Om svenskans accenter, melodi och uttal*. Studentlitteratur.

Brunelle, M. & Finkeldey, J. (2011). Tone perception in Sgaw Karen. In *Proceedings ICPhS 2011*, Hong Kong, 17-21 August 2011. (pp. 372-375).

Jones, R.B.Jr. (1961). *Karen linguistic studies: Description, comparison, and text*. Berkeley: University of California Press.

Manson, K. (2011). The subgrouping of Karen. http://www.academia.edu/620539/The_subgrouping_of_Karen_languages

Naw, V. (2011). *The Phonology of Dermuha and a Phonological and Lexical comparison between Dermuha, Sgaw Karen and Pwo Karen*. Thesis. Payap University, Chiang Mai, Thailand.

Praat: http//www.fon.hum.uva.nl/praat

Riad, T. (2013). *The phonology of Swedish*. Oxford University Press.

McCrocklin, S. (2014). Dictation programs for pronunciation learner empowerment. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 30-39). Ames, IA: Iowa State University.

DICTATION PROGRAMS FOR PRONUNCIATION LEARNER EMPOWERMENT

Shannon McCrocklin, University of Texas Pan American

Although autonomy, the capacity to learn independently, has been recognized as a language learning goal since Holec (1981) first applied the idea to language teaching, pronunciation teaching is an area that has been ignored by most proponents of autonomy. This article presents a piloted mixed-methods research study examining whether the use of Automatic Speech Recognition (ASR) as part of a hybrid pronunciation class can help foster learner autonomy more than traditional face-to-face instruction. Survey results indicate that the hybrid course group developed a greater sense of autonomy. Interview results suggest that the incorporation of ASR gave students a clear strategy for practice outside of class, expanding their repertoire of available practice strategies.

INTRODUCTION

Autonomy empowers students, allowing for effective language learning outside of the classroom. Pronunciation, however, has been an area mostly ignored by autonomy research and, without help, many pronunciation students may not know how to improve their pronunciation outside of a pronunciation classroom and may feel powerless to improve without the constant monitoring and feedback from a teacher. This research study examined the potential of Automatic Speech Recognition (ASR) technologies to foster autonomy in pronunciation learners in the context of a hybrid course, which combines traditional face-to-face instruction with ASR work days.

Autonomy

Autonomy, the capacity to learn independently, has been recognized as a language learning goal since the early work of Holec (1981), who defined autonomy as "the ability to take charge of one's learning" (1981, p. 3). Research indicates that there are educational benefits of autonomy in the learning it enables. Learner autonomy is seen by many "as a means to the end of more effective language learning" (Benson and Voller, 1997, p. 13). Autonomy allows students to work more effectively on their own, allowing them to make progress not dependent on a teacher for constant instruction and direction. Autonomous learners have also been found to have higher motivation and higher learning achievement. Classes that work to foster autonomy have been found to positively affect students' motivation and achievement (Dickinson, 1995; Furtak & Kunter, 2012; Murray, 1999). Deci and Ryan (1985) explain this relationship by claiming that students who self-determine all or some of the learning content or methods are more likely to be driven by intrinsic motivation, which stems from an interest in the task itself.

Although autonomy has potential benefits, students raised in more traditional teacher-led classrooms may devalue autonomy, appreciating more teacher-led (spoon-fed) approaches (Ming & Alias, 2007). These students may also feel uncomfortable with the idea of directing their own

learning (Luke, 2006). These students may need help and guidance to develop their autonomous learning ability.

Fostering Autonomy

To develop autonomy, this research study takes a gradualist approach to autonomy as set forth by Allford (2007). In gradualist positions, autonomy is seen as a long-term goal, something to be developed eventually. Skill in autonomous language learning, as well as proficiency and skill in the L2, is developed through study and practice. The teacher is also often considered to play a significant role in this process, providing training and guidance (Allford, 2007, p. 14).

Schwienhorst (2008), one researcher that takes a gradualist approach, suggests that fostering autonomy revolves around and depends on experimentation. Students need to be able to explore and experiment with the language itself as well as with language learning styles and strategies (p. 9). In Schwienhorst's framework, students need to be given information about learning styles and strategies as well as be guided to use these strategies and tools for experimentation with the language.

One way of doing this may be to develop a hybrid course design in which students would meet with the instructor for part of the class-time, being introduced to important language features, and then for the other half of the class-time with students doing would practice guided experimentation with the language features. A hybrid plan, however, could be complicated to apply to pronunciation teaching.

Traditional pronunciation teaching does not lend itself easily to autonomous learning or experimentation. Many pronunciation classroom activities still rely on the teacher to model "correct" pronunciation and to monitor, evaluate, and give feedback on student production. Having students monitor their own pronunciation proves difficult because students lack aural discrimination categories appropriate to the L2. Research has indicated that, for most language learners, sounds in an L2 are filtered through the phonological system of the L1 (Beddor & Strange, 1982; Blankenship, 1991; Flege, Munro, & Fox, 1993). Filtering through the L1 can lead an L2 learner to make distinctions that are inappropriate for the L2 and may prevent learners from identifying pronunciation errors when they make them. Without the ability to effectively monitor their production, students will not be able to learn how to control their motor functions to create the appropriate sounds.

While there is potential for pronunciation practice and learning to be autonomous, the task may be daunting or overwhelming to students, especially those not familiar with strategies that can help them. Students need tools that can empower them to experiment with pronunciation, without relying on the teacher for constant monitoring and feedback, tools that will help students become more autonomous as pronunciation learners.

Technology for Fostering Pronunciation Autonomy

Technology offers many tools to potentially help students work on their pronunciation. One technology that shows great promise for pronunciation experimentation work that would allow

both experimentation with the language as well as feedback is Automatic Speech Recognition (ASR), which would allow students to experiment with the language in a safe, private setting. "Automatic speech recognition (ASR) is an independent, machine-based process of decoding and transcribing oral speech. A typical ASR system receives acoustic input from the speaker through a microphone, analyzes it using some pattern, model or algorithm, and produces an output, usually in the form of a text" (Levis & Suvorov, 2012, p. 1).

When used for pronunciation training, ASR is a tool that allows students to practice at their own speed, getting feedback from the words recognized. Research has shown that ASR seems to facilitate pronunciation improvement for diverse populations of learners (Hincks, 2003; Neri, Mich, Gerosa, and Giuliani, 2008; Neri, Cucchiarini, and Strik, 2006). These studies focused, however, on student improvement, measuring accuracy gains with a pre- and post-test design. These studies did not focus on developing student autonomy and made no effort to measure changes in autonomy.

There seems to be little overlap between research in autonomy and research into pronunciation. This research study seeks to bring these two fields together, with the goal of using ASR to foster autonomy in pronunciation learners.

Research Questions

This study seeks to answer the following questions. Does the utilization of technology tools for pronunciation feedback in a hybrid course:

foster learner autonomy?

lead to increased beliefs of empowerment to improve pronunciation ability?

Methods

To answer these questions, a mixed methods approach was taken to measure changes in stated beliefs of autonomy and empowerment over time with surveys (quantitative) and to better understand participants' beliefs through interviews (qualitative). Each participant completed a survey and interview before and after participating in a three-week pronunciation workshop covering vowels and consonants known to be problematic for many ESL learners.

Participants

Participants were recruited from an undergraduate ESL writing course. Seven participants completed the research study, three males and four females. Six out of the seven participants were native Chinese speakers; one was a native Portuguese speaker.

The surveys and interviews

The stated beliefs of the participants' autonomy and feelings of empowerment were measured at the beginning and the end of the course through surveys and interviews. The surveys were used to determine if the attitudes and skills before and after the class were different. The surveys were administered through SurveyMonkey.com. The pre-course survey had 10 questions that addressed participant background information, language learning habits, as well as nine Likert scale question items that addressed participants' beliefs about their autonomy and empowerment. The post-course survey asked participants to answer the same Likert scale items used in the pre-course survey.

The interviews were primarily used to identify the student's perceptions of the causes of changes in survey responses. Most of the post-interview questions aimed to elicit the student's repertoire of tools, skills, or strategies the students had developed during the course.

The Pronunciation Workshops

Participants in the study were asked to take a three week pronunciation workshop. Volunteers for the study/workshop were semi-randomly assigned to one of two groups (based on times available and technology owned):

Control (traditional face-to-face course)

Experimental (hybrid with technology/online day)

Both groups participated in workshops with two work days per week and one homework assignment. For both courses, the first workday met together as a class doing listening practice with the sounds, controlled production activities, and guided production activities. Students were also introduced to spelling patterns for the targeted vowel sounds. Finally, both courses were introduced to pronunciation practice strategies, including focused listening, practicing with ASR (through Windows Speech Recognition or voice search on smart phones), and covert rehearsal, a form of private practice in which learners monitor their speech for pronunciation issues or errors (Dickerson, 1994).

The second workday included a listening review, but focused mostly on production for both class, but differed in format. For the traditional course the second workday was again face-to-face and instructor led, but the hybrid course moved online using technological tools. For the hybrid this included listening activities done through online listening activities and production practice performed with software already a part of Windows, Windows Speech Recognition (WSR). Students were directed to monitor the dictation provided by WSR and work on correcting their pronunciation if the program was not able to correctly identify targeted sounds in their intended word. The computer based listening and production activities for the hybrid course were managed through Moodle, a course management website.

Both classes were asked to submit a recorded file each week as homework. The assignment asked participants to record activities that demonstrated work in these areas. Figure 1 shows how the two courses aligned for each week:

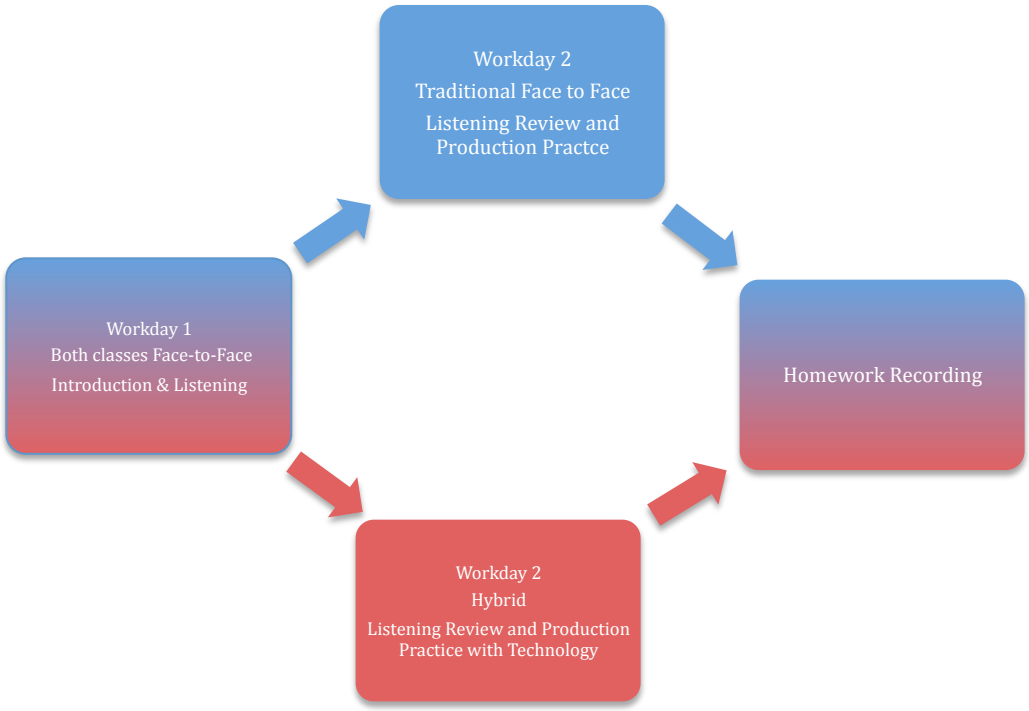


***Figure 1*. Course Design for each Week**

The topics for the course, the vowel pairs /ɛ/ vs. /æ/, /a/ vs. /ʌ/, and /i/ vs. /ɪ/ and the consonants /ɹ/, /θ/, /ʒ/, and /dʒ/were chosen mainly based on the likelihood of the sounds being a problem for the participants, using Nilsen and Nilsen (2002) to identify the contrasts that are likely to be problematic to speakers of many languages. Because this course was short, it was neither possible to make a comprehensive class covering all of the contrasts problematic for all students, nor to design a course in advance without already enrolled participants that could target problems for the particular class. It is important, however, that students find as much of the training valuable as possible so it is important that students recognize a need for the training. This is more likely to happen if students do not have mastery of a sound or contrast. Thus, while functional load as proposed by Brown (1988) was considered, I decided that it was more important to choose sounds that were likely to affect most students.

Results

Survey

Results from the surveys indicate that generally both courses seem to have increased beliefs of autonomy and empowerment. The averages for both groups before and after the course on each item are displayed in Table 2, along with the calculated difference between the groups.

Table 1
Average Group Responses by Item

	F2F Pre	F2F Post	**F2F Diff**	Hy-brid Pre	Hy-brid Post	**Hy-brid diff**
I am concerned about my accent in English.	4.33	4.67	**0.34**	5.5	5.25	**-0.25**
I feel that I practice my pronunciation in English frequently.	3.33	4.0	**0.67**	3.5	4.75	**1.25**
I can hear when I mispronounce a sound or word in English.	4.0	4.67	**0.67**	3.5	4.5	**1.0**
I am aware of different ways to practice my English pronunciation.	3.33	5	**1.67**	5.25	5.75	**0.5**
I am prepared to practice my English pronunciation on my own.	4.33	4.33	**0**	4	5	**1.0**
I feel like I need to hear a native speaker to know how to produce a word.	5	4.67	**-0.33**	5.25	4.25	**-1.0**
I have tools that can help me work on my pronunciation.	3.67	4.67	**1**	4.5	5	**0.5**
I feel like I need a native speaker of English to correct me on my pronunciation to improve.	5	4.67	**-0.33**	5.5	4.5	**-1.0**
I feel that I have the ability to improve my English pronunciation on my own.	4.33	4.0	**-0.33**	3.5	4.75	**1.25**

It is important to notice that in this chart, scores decline for "I feel like I need to hear a native speaker to know how to produce a word" and "I feel like I need a native speaker of English to correct me on my pronunciation to improve." Lower scores on these items suggest higher autonomy and empowerment because they suggest less reliance on others. To give a more accurate representation of the changes, Figure 2 and 3 separate out the "anti-autonomy items", those for which ***lower*** scores are thought to indicate higher autonomy, and "pro-autonomy items", those for which a ***higher*** scores are thought to indicate higher autonomy.

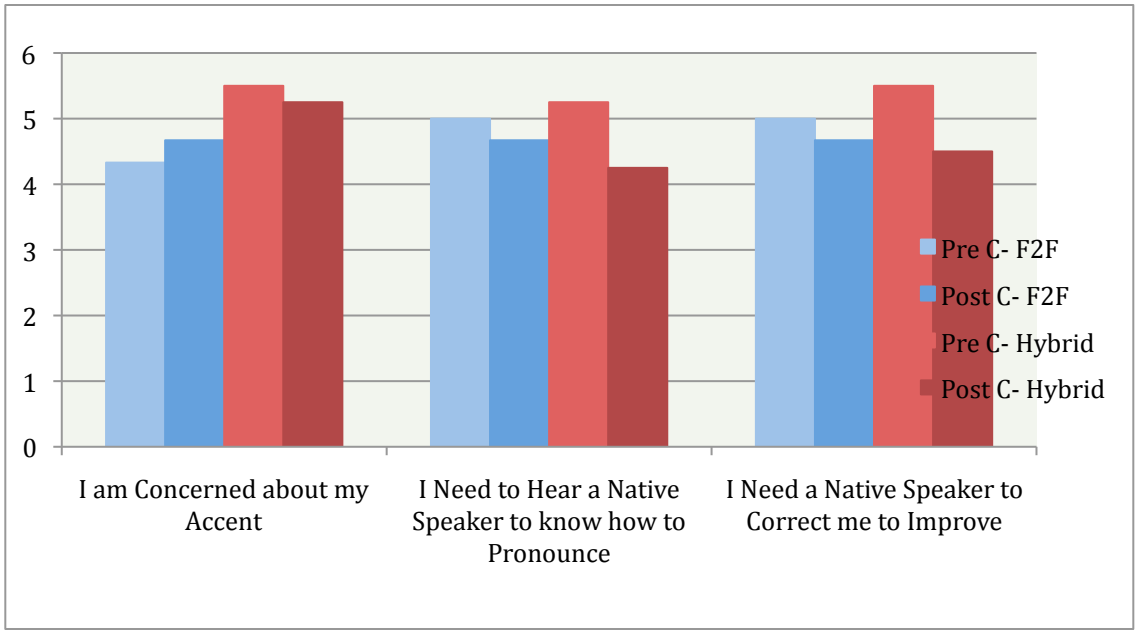


Figure 2. Anti-Autonomy Items by Group

While the Face-to-Face group grew more concerned about their pronunciation (Figure 2), the Hybrid group became less concerned about their pronunciation. Also, the hybrid group had a more decreased reliance on native speakers than the face-to-face group for both hearing and correction.

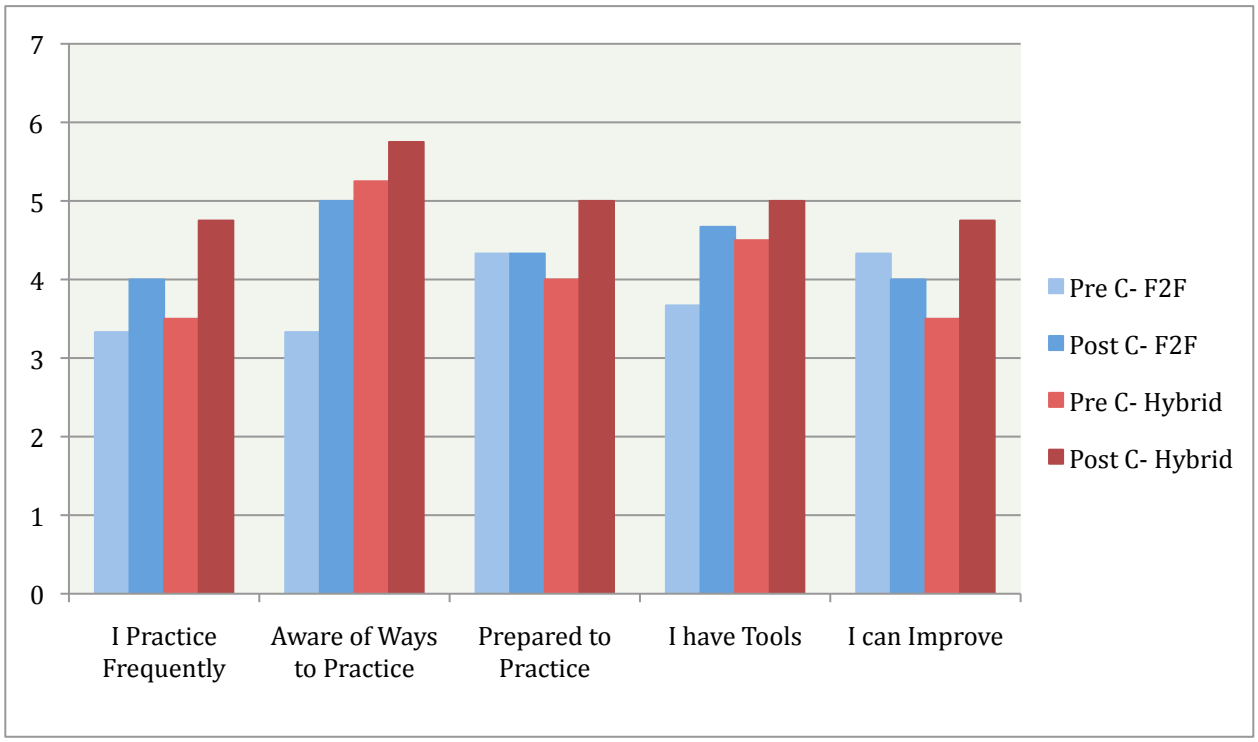


Figure 3. Pro-Autonomy Items by Group

In Figure 3, the hybrid group had a greater difference (more gains in autonomy) on three items "I feel that I practice my pronunciation in English frequently, " "I am prepared to practice my English pronunciation on my own," and "I feel that I have the ability to improve my English pronunciation on my own." It is surprising to note, though, that the hybrid group had lesser gains on two items "I have tools that can help me work on my pronunciation" and "I am aware of different ways to practice my English pronunciation." This might have been partially due to the fact that the hybrid group started with higher levels of agreement with those items (they had less room to "improve"). The hybrid group did have higher levels of agreement on those items, though, at the end than the face-to-face group did, with a score of 5.75 out of 6.0 on the prepared to practice item and a 5.0 out of 6.0 on the tools question.

Interviews

The pre- and post- workshop interviews were focused on the participants' language learning/pronunciation practice skills. In both the pre- and post-interviews, questions were asked what tools or skills participants used as part of their pronunciation practice repertoire. Before the workshop, most participants reported watching movies and using dictionaries. When asked what they did while watching movies, participants responded that they just listened to the language, hoping it would help.

After the course, both groups had expanded their repertoires. One member of each group mentioned that the spelling rules were of great help, and they were working to practice and apply those patterns. Also, one member of the face-to-face group mentioned using covert rehearsal, the one non-technology based strategy introduced during the course. Both groups were introduced to the idea of using voice recognition software as a pronunciation practice strategy, but none of the face-to-face participants mentioned using voice recognition. On the other hand, two members of the hybrid course specifically mentioned enjoying using Windows Speech Recognition and their plan to continue doing so. A third member of the hybrid group also mentioned continuing the activities done as homework, which would include work with WSR, but did not specifically mention WSR.

It is also interesting to note that while both groups still mentioned movie or video watching as part of their plan for future work, two of the hybrid group specifically mentioned enjoying or planning to continue work with TED talks, which were introduced in the course. One of these members also described the value of focused listening that students worked with in the course.

DISCUSSION

While this study included a limited number of students, it showed promising results that guided work with technological tools may help students feel more autonomous and empowered in their language learning ability, specifically in regards to pronunciation. Results from the survey seem to suggest that the hybrid group did develop their sense of autonomy and empowerment more than the face-to-face course did. Although both groups were introduced to the same pronunciation practice strategies, the hybrid group, which was asked to spend time working alone with the technologies and strategies seems to have increased their sense of autonomy more.

Although there were two items in which the hybrid group's autonomy beliefs scores increased less, this might have been due to the high starting scores. The hybrid groups did end, though, with higher scores (post-course) on those items than the face-to-face group.

Comments from the interviews suggest that guided practice with Windows Speech Recognition led to an expansion of students' repertoire of language learning/pronunciation practice strategies for pronunciation work. WSR was mentioned in student's plans for continued pronunciation work and allowed for more specific plans for continued learning.

These findings support the idea that use of technology, which can provide an opportunity for experimentation (Schwienhorst, 2008) in a safe environment (Banafa, 2008), can enhance student autonomy (Benson, 2011) for language learners, including fostering autonomy for pronunciation learning, an area that has traditionally been very teacher dependent.

ABOUT THE AUTHOR

Shannon McCrocklin is an assistant professor at the University of Texas-Pan American. She earned her Ph.D. in Applied Linguistics and Technology at Iowa State University in Ames, IA. She holds an M.A. in Teaching English as a Second Language from the University of Illinois at Urbana-Champaign where she developed an interest in pronunciation teaching and applied phonetics and phonology. Shannon has taught English pronunciation to undergraduate and graduate students as well as to international faculty at Iowa State and the University of Illinois. Her research focuses on improving pronunciation training for students and CAPT (Computer-Assisted Pronunciation Teaching). She has presented at CALICO, NCTE, PSLLT, and AAAL.

REFERENCES

Allford, D. (2007) Introduction: language, autonomy, and the new Learning Environments. In D. Allford & N. Pachler (Eds.), *Language, Autonomy, and the New Learning Environments* (pp. 11-34). Bern, Switzerland: Peter Lang.

Banafa, F.H. (2008) *Effects of IT on Pronunciation*. LaVergne, TN: Lightning Source Inc.

Beddor, P.S. & Strange, W. (1982). Cross-language study of perception of the oral-nasal distinction. *Journal of the Acoustical Society of America*, *71*, 1551-1561.

Blankenship, B. (1991). Second language vowel perception. *Journal of the Acoustical Society of America*, *90*, 2252-2252.

Benson, P. (2011) *Teaching and Researching Autonomy* (2nd ed). Harlow: Pearson Education Limited.

Benson, P. and Voller, P. (1997) Introduction: autonomy and independence in language learning. In P. Benson & P. Voller (Eds.), *Autonomy and Independence in Language Learning* (p. 1-12). Edinburgh: Addison Wesley Longman Ltd.

Brown, A. (1988) Functional Load and the Teaching of Pronunciation. *TESOL Quarterly*. *22(4)* p. 593-606.

Deci, E. L. & Ryan, R. M. (1985) *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum Press.

Dickerson, W. B. (1994). Empowering students with predictive skills. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New views, new directions.* Alexandria, VA: TESOL, Inc.

Dickinson, L. (1995) Autonomy and motivation: a literature review. *System, 23(2),* 165-174.

Flege, J.E., Munro, M.J., & Fox, R.A. (1993). Auditory and categorical affects on cross-language vowel perception. *Journal of the Acoustical Society of America*, *95*, 3623-3641.

Furtak, E.M. & Kunter, M. (2012) Effects of autonomy-supportive teaching on student learning and motivation. *The Journal of Experimental Education*, *80*, 284-316.

Hincks, R. (2003) Speech technologies for pronunciation feedback and evaluation. *ReCALL*. *15(1),* 3-20.

Holec, H. (1981) *Autonomy in Foreign Language Learning.* Oxford: Pergamon.

Levis, J. & Suvorov, R. (2014). Automated speech recognition. In C. Chapelle (Ed.) *The encyclopidia of applied linguistics*. Retrieved from http://onlinelibrary.wiley.com/store/10.1002/9781405198431.wbeal0066/asset/wbeal0066.pdf?v=1&t=htq1z7hp&s=139a3d9f48261a7218270113d3833da39a187e74.

Luke, C. (2006) Fostering learner autonomy in a technology-enhanced, inquiry-based foreign language classroom. *Foreign Language Annals. 39(1),* 71-86.

Ming, T.S., & Alias, A. (2007) Investigating readiness for autonomy: A comparison of Malaysian ESL undergraduates of three public universities. *Reflections on English Language Teaching, 6(1),* 1-18.

Murray, G. (1999) Autonomy, technology, and language-learning in a sheltered ESL immersion program. *TESL Canada Journal*, *17(1)*, 1-15.

Neri, A., Cucchiarini, C., & Strik H. (2006) ASR-based corrective feedback on pronunciation: does it really work? *Proceedings of the ISCA Interspeech 2006*, Pittsburgh, PA, 1982–1985.

Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008) The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, *21(5)*, 393-408.

Nilsen, D. L. & Nilsen, A.P. (2002) *Pronunciation Contrasts in English*. Prospect Heights, IL: Waveland Press.

Schwienhorst, K. (2008) *Learner Autonomy and CALL Environments*. New York: Routledge.

Cauldwell, R. (2014). Listening and pronunciation need separate models of speech. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 40-44). Ames, IA: Iowa State University.

LISTENING AND PRONUNCIATION NEED SEPARATE MODELS OF SPEECH

Richard Cauldwell, Speech in Action, Birmingham UK

Language learning pedagogy is currently dominated by the Careful Speech Model (CSM), key components of which are the citation forms of words, the relationship to grammatical categories (tone groups to clauses, intonation to questions, etc) and the rules of connected speech. However these aspects of the CSM – and particularly the rules of connected speech – are inadequate to characterize the unruliness, wildness, and massive reductions that occur in spontaneous speech. To teach listening more effectively we need a new model of speech – the Spontaneous Speech Model (SSM) – some of the components of which are outlined in this paper. I also argue: that we language teachers need a new mind-set which permits and accepts conflicts between what we teach in the CSM and what we teach in the SSM; that we need to take active steps to guard against our expert-listener tendency to hear full words where only massively reduced traces occur; that we need to let go of some of our favorite rules of connected speech.

The learning ghetto

A friend of mine – Tim – learnt Welsh as an adult. He told me of his pride in being able to have a conversation with his teacher in a non-classroom setting. But this pride turned to dismay when they were joined by a native speaker who then conversed with his teacher in spontaneous speech that Tim could not understand. From being confident and comfortable in Welsh, he suddenly felt excluded. He described this experience as feeling he was trapped in a 'language-learning ghetto'.

What Tim has identified is the fact that we ill-prepare language learners to handle the stream of spontaneous speech in real time. I believe that there is increasing realisation in the language teaching profession of the fact that we are not quite doing the right thing with the teaching of listening. Rost (2001: 13) writes that listening '…is still often considered a mysterious *black box* for which the best approach seems to be *more practice*...' We have yet to work systematically to open up the black box, and to go beyond the more-practice paradigm.

Separate goals

The way forward is pointed by Celce-Mucia et al (2010, p. 370) who state that the speech learners need to understand is 'much more varied and unpredictable than what they need to produce in order to be intelligible'. They conclude that for pronunciation and listening, 'the goals for mastery are different'.

My contention is that because the goals for mastery of pronunciation and listening are different, **we need two different models of speech –** one for pronunciation, which we already have, and an additional one (which we need to drag into existence) for listening. This model needs to be one that is an accurate representation of the realities of spontaneous speech. The trouble is that ELT is dominated by the model we already have, the Careful Speech Model in its British and North American variants (Cauldwell, 2013). The Careful Speech Model (henceforth CSM) is based on the citation forms of words, and consideration of how these citation forms would sound when joined together and placed in clauses and sentences. Although it is useful for teaching pronunciation, this model of speech is not appropriate for teaching listening to everyday speech.

Need for a new model

In order to teach effectively towards a particular goal, we need to have a descriptive model of what we want learners to master, from which we can derive a syllabus, a metalanguage and teaching/learning activities. The CSM does not give us a model of spontaneous speech, it is a

separately potted in a greenhouse (isolated citation forms), and of flower arrangements in a garden (rules of connected speech). However, for the purposes of teaching listening we need to prepare learners for the **jungle** of spontaneous speech, where the vegetation is crushed together in a messy and unruly manner which is quite unlike any orderly garden arrangement. Crucially, the mess and unruliness of spontaneous speech go way beyond what is generally allowed for in presentations of the rules of connected speech (such as those described in Brown, 2010; Celce Murcia et al, 2010; Roach, 2009).

So we need to embrace the mess and unruliness of everyday speech. And we will (for the purposes of listening) have to be prepared to let go of some of our favorite things to teach, to devise another model of speech, the Spontaneous Speech Model (henceforth SSM) which is a better description of the Jungle.

Many of the components of the CSM are expressed in terms of rules, and are exemplified with scripted acted speech. But acted speech is a completely different animal from spontaneous speech. The rules of the CSM can be viewed as hypotheses about what happens in spontaneous speech. But I suspect that most of them either have been refuted (as has happened for stress timing cf. Dauer, 1983, Roach, 1982; and for emotion in speech cf. Stibbard, 2001) or will be refuted. It is my belief that in spontaneous speech, the null hypothesis will be found to reign supreme. We should not expect that any of the CSM rules are true of spontaneous speech.

So one of the features of the spontaneous speech model (SSM) should be that it is unhooked from the expectations and rules of the CSM. I am not suggesting that we do away with the CSM – we should continue to allow it to perform the role that it has done for a long time, which is to act as a syllabus and a provider of metalanguage for the teaching of pronunciation.

Suggestions for components of the new model

There have been a number of scholars who have over the years pointed us in the direction of an SSM. The two editions of Brown's work (1977; 1990) and Shockey (2003) on the processes that operate on natural speech are landmarks, and in the mid nineteen-eighties Ur (1984, p. 17) suggested that we need to focus on words which are spoken ‘quickly, in an unemphasized position in a sentence and juxtaposed with other words ...'. More recently, Field (2008, p. 196) suggests that the unit of decoding and understanding is neither the word nor the sentence, it is the intonation group: ‘Words take their shapes from the intonation group as a whole and may not be identifiable until the whole group has been heard’.

A number of authors have favorite examples which they quote to show the extreme reductions that can happen in spontaneous speech. For example, Field (2003, p. 331) gives us ‘narp meme’ which is a reduced form of ‘Know what I mean?’. And a number of the contributors to Brown (2010) present examples of extreme reductions embedded in their classroom activities. Their examples point to a jungle-like speech world which lies beyond the rules of connected speech. These contributors include Jeff Stewart (ibid. p. 9), Michael Wilkins (p. 25), James Cassidy (p. 94), Mitsuyo Nakano (p. 224) and Arthur Nakano (p. 248). It is important that we do not regard such examples as interesting oddities, as exceptions to the rules which are familiar to us. They need to be put center stage, and treated as typical examples of what can and does happen in spontaneous speech: they need to be regarded as central to the model, not as outlying oddities. In phonetics and psycholinguistics there is an emerging literature on extreme reductions in which Johnson (2004) and Ernestus & Warner (2011) are key papers. This literature can provide important input to this model.

Escape from the graphic

We need also to escape the temptations and expectations of the forms of the written language, the graphic substance, which is the means by which most language teaching and learning takes place. We have to come to terms with the idea that students need to master the transient sound substance

that is invisible and shaped by the speaker in real time. It is not a permanent graphic substance that is visible, it is not shaped by rule from the citation forms and the grammar.
The graphic substance can be seen, and it remains in view so that we can inspect it and edit it. And although there is some variation in the shape of words (different fonts, upper and lower case letters, some contractions) the words retain a close resemblance to their canonical shape.
Sound substance, on the other hand, is invisible. When it has happened it is gone – leaving traces in short-term memory which are then replaced by the traces for the new speech which is happening. And words (most often) do not occur on their own: they occur in rhythmic chunks – speech units – with some words emphasised and others not. And words are run together in a continuous acoustic blur where it is difficult to determine where one word ends and another begins. For learners the stream of speech is a continuously changing uncapturable fast-moving mush of sound substance.

A new mindset
In order to help learners master this sound substance we need to change some of our ideas and practices. So when we are working on the goal of listening, we need to accept that some of the things that we need to do may well be in conflict with our practices in the teaching of pronunciation, and in conflict with the CSM. This is because many of the rules and guidelines of the CSM are violated in spontaneous speech.

Drafting phenomena
Perhaps the most noticeable violation is the presence of drafting phenomena. In the CSM we advocate a tidy version of the language, akin to carefully edited written language. But spontaneous speech contains – as part of the sound substance – phenomena which would be edited out and would therefore be invisible in the written language. Because it is created and heard in real time, we hear the drafting phenomena: pauses for thought, filled pauses, references to speaker and hearer roles (‘I mean’, ‘you know’), references to encoding difficulties (‘How shall I put it’), appeals to shared understanding (‘you know what I mean’) and the changes of mind. These phenomena are crossed out in the written draft, and do not appear in the final version of a written text. But they are heard in the sound substance. In spontaneous speech, unlike in writing, you cannot stop for thought, erase or edit what you are saying without contributing to the sound substance – even if that contribution is silence. The CSM deprecates such phenomena (and things such as ‘like’, ‘and that kind of thing’ ‘and stuff’) but this stuff occurs in spontaneous speech, and if we are to teach the sound substance of everyday speech, we should not pretend such phenomena don’t exist. They can be kept out of the CSM, but they definitely need to be in the SSM.

Sound shapes and the blur gap
We need to make a deliberate effort to put the CSM out of our minds when we are teaching listening. This is because the CSM primes us to attend to spontaneous speech in a way which conforms to our expectations: its rules and guidelines become expectations, and that’s what we hear. This leads to something I refer to as the blur gap (Cauldwell 2013, p. 17) which is the gap between what expert listeners believe they hear, and what is actually present in the sound substance. Countering the effects of the blur gap is not an easy thing to do. This is because, as an expert native speaker listener, I construct much of what I believe I hear. The sound substance which arrives at my ears consists of traces of words streamed in the mush of the acoustic blur of spontaneous speech. But I am largely unaware that these 'words' are made up of mushed blurred traces which can be very distant from the citation form. I will generally perceive a word-trace as something more substantial, as having a fuller soundshape closer to the citation form. Our ears play tricks on us. Our expertise with the CSM plays tricks on us. We do not have a good sense of what is happening in the stream of speech, so we need the help of technology to train our ears. One way of doing this is to use

Audacity to cut up the stream of speech (Cauldwell 2013, p. 296) and to create audio concordances of words in speech units.

Letting go

As well as guarding against the blur gap, we need to let go of our attachment to the rules underlying (famous) examples of how speech works. For example word pairs such as 'might rain' and 'my train' which the CSM holds to sound different, may actually sound identical. Equally, different tenses may sound identical: the present tense 'they bought' might well sound identical to the past perfect 'they'd bought'. The underlying articulations may reflect the differences in grammar, but what is present in the sound substance may be identical. Thus distinctions which are very important in in the CSM may well be indistinguishable, or indeterminate, in spontaneous speech. This indeterminacy needs to be built in to the SSM. This particularly applies to negation – negative morphemes may be inaudible, so that there is no difference, in the sound substance between 'very legal' and 'very illegal'.

Conclusion

The late Professor John Sinclair of the University of Birmingham (UK) used to tell an anecdote about a German professor of English who complained to him that 'You teach us one English, and then mumble to each other in another English that we don't understand'. Like Tim, with Welsh, the German professor feels excluded from interactions s/he wants to be part of. We need to improve the teaching of listening by adopting a new goal, a new model (the SSM), and changing our mindset so that we don't exclude our learners from interactions that they really want to be part of.

ABOUT THE AUTHOR

Richard Cauldwell has taught English in France, Hong Kong, Japan, and the UK, where he worked at the University of Birmingham. Since 2001 he has published electronic and print materials for listening and pronunciation. Two of his publications have won British Council Innovations in ELT awards (ELTons): *Streaming Speech* (online course) in 2004 and *Cool Speech* (iPad app) in 2013. His *Phonology for Listening: Teaching the Stream of Speech* published in print in 2013 and as an eBook in 2014 has been shortlisted for three awards, and has received high praise in reviews from many sources, including *TESOL Quarterly*, and the *Journal of the International Phonetic Association*.

REFERENCES

Brown, G. (1990). *Listening to spoken English,* 2nd ed. Cambridge: Cambridge University Press.

Brown, J. D., Ed. (2012). *New ways in teaching connected speech.* Alexandria, VA: TESOL International Association.

Cauldwell, R.T. (2013). *Phonology for listening: Teaching the stream of speech.* Birmingham: Speech in Action.

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

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics, 11*, 51-62.

Ernestus, M., & Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics, 39*. 253-260.

Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal, 57*(4), 325-334.

Field, J. (2008). *Listening in the language classroom*. Cambridge: Cambridge University Press.

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama, & K. Maekawa (Eds.), *Spontaneous speech: Data and analysis* (pp. 29-54). Tokyo: The National International Institute for Japanese Language.

Roach, P. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. In D. Crystal (Ed.) *Linguistic Controversies, Essays in Linguistic Theory and Practice,* 73-79. London: Edward Arnold.

Roach, P. (2009). *English phonetics and phonology: A practical course,*4th ed. Cambridge: Cambridge University Press.

Rost, M. (2001). *Listening.* In R. Carter, and D. Nunan (Eds.) *The Cambridge guide to teaching English to speakers of other languages* (pp. 7-13). Cambridge: Cambridge University Press.

Shockey, L. (2003). *Sound patterns of spoken English*. Oxford: Blackwell.

Stibbard, R. (2001). *Vocal expression of emotions in non-laboratory speech: An investigation of the Reading/Leeds emotion in speech project annotation data*. Doctoral dissertation: University of Reading.

Ur, P. (1984). *Teaching listening comprehension.* Cambridge: Cambridge University Press.

Violin-Wigent, A. (2014). Comparing online vs. face-to-face classes: A case study of a a French pronunciation class. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 45-57). Ames, IA: Iowa State University.

COMPARING ONLINE VS. FACE-TO-FACE CLASSES: A CASE STUDY OF A FRENCH PRONUNCIATION CLASS

Anne Violin-Wigent, Michigan State University

This study compares the outcome of a traditional face-to-face (F2F) class, taught in the spring of 2012, with its online equivalent, taught in the summer of 2013. Based on comparable assignments, results show that there are mostly no significant differences on written and oral quizzes (except for two oral quizzes) and that the evolution of students is similar in both formats. The lower performance of the online class on these two assignments may be due to the difficulty of the elements included. In these cases, the presence of the instructor as a motivator and a source of immediate feedback seems to be beneficial. In one other oral quiz though, the F2F class performed at the same level as the online class even though it is also a fairly difficult theme (liaisons). It seems, for this assignment, that the option to control one's learning, possible in the online class, was more beneficial than the immediate feedback from the instructor,.

INTRODUCTION

In the past ten years, all levels of education have seen a drastic increase in online classes and learning. A cursory look in The Chronicle of Higher Education gives a plethora of examples showing that, even several years ago, enrollment in online college-level classes is growing much faster than general enrollment in college:

> "From the fall of 2004 to the fall of 2005, the enrollment [in online classes] grew by 36.5 percent. And from the fall of 2005 to the fall of 2006, enrollment increased 9.7 percent. Still, it's growing faster than general college enrollment, which grew by 1.3 percent from 2005 to 2006" (Carnevale 2007).

One of the main consequences of this explosion was a call for quality control and learning assessment. Indeed, in 2012, a supplement of The Chronicle of Higher Education, entitled "Taking Measure of Online Education" was entirely dedicated to this very issue: http://chronicle.com/article/Online-Learning-Supplement-/131624/. Several studies have shown that technology can help students learn language by putting them in control of their learning, by enabling them to collaborate in new ways, and by exposing them to more authentic language (Blake, 2008; Goertler & Winke 2008a; Lancashire 2009). Additionally, they can study in a time frame that is most convenient for them.

The current study adds to this discussion of the effectiveness of online language classes. In particular, it focuses on a phonetics and pronunciation class when most studies focus on general language classes, on grammar, or on writing. The primary goal of this study is to measure whether students in an online class obtained similar results on written and oral

quizzes as those enrolled in a traditional face-to-face (henceforth F2F) class, asking the following research questions:

1. Are the grades on assignments comparable in the F2F class and online class?
2. Is the evolution of the pronunciation of students between the first recording and last one comparable?

Previous studies on online learning

Before delving into the study, it is important to establish definitions. A hybrid class is generally defined as a class containing at least 50% of the content online whereas a class with less than 50% of online content is typically referred to as technology-enhanced. A class with 100% of the content online and with no actual in-person contact is called an online class.

Attempting to review all the studies published on the topic of online learning is well beyond the scope of the current paper. I will only briefly mention some studies (among too many to list here) that relate most to the effectiveness of online learning of a foreign or second language.

Overall effectiveness

Many studies show that learning can and does take place in an online environment, even though Goertler and Winke (2008b) report that many experiments (successful or not) go unreported. They urge more publications of results so that "others can follow in their footsteps but avoid making some of the same mistakes" (Goertler & Winke 2008b, p. 254). Sanders (2005) compared student learning outcomes (as measured by the BYU placement exam) before and after the implementation of a redesigned curriculum that reduced the amount of face-to-face class time and replaced it with various forms of online activities. His results show comparable scores on the placement test for both instructional formats. In addition, he reported that there was no difference in student retention nor in course completion. Since the introduction of this online component allowed for significant savings (in terms of instructor hours as well as student cost), these were overall very positive results. Sagarra and Zapata (2008) reported the positive effect of an online workbook on the results on grammar tests. The 245 Spanish-language students in the study reported enjoying the usefulness and ease of access of the workbook as well as the immediate feedback it provided. Blake (2009) also reported no difference between F2F and online/hybrid classes. Abuseileek (2009), trying to study smaller elements of the overall learning process, compared the acquisition of grammar under four different treatments with computer-based or non-computer-based learning inductively or deductively. His results showed that computer-based technique seemed to facilitate the learning of more complex structures (described as complex sentence structures), especially if associated with a deductive format. Less complex structures did not show a difference across all four treatments.

A recent study, however, qualified these positive results by introducing the variable of class size. Russell and Curtis (2013) compared a small Spanish class (25 students) to a larger one (125 students), both of them taught fully online. Not surprisingly, students in the larger class were less satisfied with their experience than those in the smaller class. In the larger class, there was much less interaction, not only with the instructor, but also among students. They concluded that a larger class created an environment that was less conducive to effective learning than the smaller class. As administrators often see online classes as a money-saving device, this study provides crucial information regarding the limit of the effectiveness of online classes. This is echoed in the high school environment, as reported by Oliver, Kellog, and Patel (2012). High school students enrolled in a virtual public school reported lower satisfaction rates with their experience learning a foreign language than they did for other subjects. In particular, students felt that there was inadequate support and insufficient collaboration among students. These two studies point to one of the most often cited problem with online instruction, namely, the lack of interaction and connection.

Limited interaction

As McBride and Fägersten (2008) point out, students enrolled in online classes reported an increased feeling of isolation, because they found themselves interacting primarily with the computer. Interestingly, this also applied to instructors. A growing number of studies underscore the lack of community (White 2003 and Wildner-Bassett 2008 among many others). In an online class, there is no eye contact, no group work, no arriving to class early to chat with fellow students, and no staying after class to talk to the instructor. For this reason, McBride and Fägersten (2008) underline the importance of a steady stream of communication between instructors and students.

More recently, some have attempted to evaluate and provide solutions to remedy this important issue. Eneau and Develotte (2012) reported that graduate students enrolled in an online Master's program developed ways to connect with other students to achieve a sense of community. They add, however, that this was by no means easy to do. Senior (2010, p. 146) stressed the importance of the instructor in establishing this sense of community and connectivity:

> It has suggested that, regardless of the degree to which their conventional roles are technologically supported or sidelined, language teachers need to find ways of developing and maintaining connections with their students. To do this they need to understand the complexity and dynamic nature of their roles – and be prepared to redefine them where necessary.

This study points to the need to redefine the role of the instructor in online instruction. This conclusion is also found in one of the very few studies focused on online learning of pronunciation. Indeed, Brudermann (2010) reports that there needs to be a structure in place to help students develop a sense of responsibility in front of distance learning and that this falls onto the instructor.

METHODOLOGY

The class

To conduct the present study, a fully online class was compared to its traditional F2F equivalent. The class in question was a third-year French phonetics and pronunciation class, which the author has been teaching almost every semester since 2006. It is required for all French majors and minors at Michigan State University. The F2F class in this study was taught in the spring of 2012 (a 15-week semester) with 20 students enrolled and met twice a week for 1 hour and 20 minutes. The online class was taught in summer of 2013 (a 13-week semester) with only eight students enrolled. It was the first time the class was taught online. Both classes used the same textbook: Violin-Wigent, Miller and Grim (2013). Although the F2F class used a preliminary version of the textbook, there were no significant changes in the content between the two classes. Most students enrolled in class are not primary French major but oftentimes have selected French as their secondary major or as their minor.

In both classes, the type and sequencing of activities were similar. Each lesson starts with listening comprehension activities, which are the bases for the following rule-induction phase, followed by discrimination activities. In the F2F format, this is done at the end of class so as to prepare students to read the actual lesson at home before the next class, as well as prepare a couple of activities. This is then followed by practice activities, starting with oral practice (repeating after the instructor or audio files) and transcription activities. Students enrolled in the online class where strongly encouraged to follow this sequencing to optimize their learning. The rationale behind the sequencing of activities (pre-reading activities before reading the lesson, importance of the inductive approach, focus of oral tasks) was given to both classes. The F2F class typically contains only one or two group activities per day, during which students discuss each other's homework before correcting as a class. This was eliminated in the online class.

The tasks

The study is based on the grades students received in the class for written and oral quizzes. The F2F class was given a total of nine oral assignments but only seven are included in the study as they are the ones that are identical for both classes. These oral assignments were short authentic texts (between 93 and 130 words) that students were asked to read aloud and record using Audacity. Though 20 students were enrolled in the F2F class, the results presented in Table 3 below do not always contain 20 students, as several assignments were not turned in. By comparison, all students completed all the oral assignments in the online class. In both classes, oral quizzes were assigned after extensive practice that provided students with models and feedback. In the F2F class, feedback was immediate after all students repeated after the instructor in class, whereas the online class received delayed feedback after oral homework but individualized feedback as the instructor provided a personal list of problems to each student.

Three written tests were included in the study. These are obviously different written tests for both conditions to prevent cheating. However, they all followed the same format

starting with a phonetic discrimination asking students to recognize the sound they heard, followed by theoretical questions (such as definitions, identifications, etc.), and ending with IPA transcription. For the first test, the transcription section involved reading in the IPA and writing using the normal orthographic alphabet whereas for the next two tests, students had to transcribe a series of short sentences using the IPA.

Statistical analysis

To analyze the results, a series of t-tests were run on the tasks described above . Results below do not reflect the Bonferroni adjustment even though, technically, it should be included when one runs a series of t-tests. The Bonferroni adjustment would dictate that the p values would have to be under .007 to show a statistical difference between the F2F and online classes. Since none of the p-values are that low, it was decided not to include the Bonferroni adjustment in order to see some tendencies, with the understanding that analyses and conclusions will be tentative and prudent.

RESULTS

Comparison of overall results

Table 1 below shows how many students received what final grade in each of the classes. As a reminder, there were 20 students in the F2F class and eight in the online class.

Table 1
Overall grade distribution

	F2F		Online	
2 (C)	2	10%	1	12.5%
2.5 (C+)	5	25%	1	12.5%
3 (B)	2	10%	1	12.5%
3.5 (B+)	3	15%	2	25%
4 (A)	8	40%	3	37.5%

Table 1 shows a similar spread in the fact that no one earned less that than a grade of 2 or C. Additionally, the proportion of students earning this lowest grade is comparable in both format, just like the proportion of students who earned the highest grade (4 or A). The major difference between the two classes can be seen in the distribution of students in the middle grades. As can be seen, there is a higher percentage of 3.5 (B+) in the online format and a higher percentage of 2.5 (C+) in the F2F class. Since there is such a small number of students in the online class, it is difficult to come to definite conclusions. We can, however, hypothesize that the online format may have attracted more conscientious students or students who are better at time management and that the results are not due to the format itself but rather to a self-selected population. Additionally, since the online class was taught in the summer, it could be the case that students were not enrolled in other classes and, therefore, had more mental energy to devote to the class.

Comparison of 3 written quizzes

In order to answer the first research question, we will first look at the results of written quizzes before moving on to oral quizzes.

Table 2
Comparison of written quizzes

	F2F Mean (SD), n = 20	Online Mean (SD), n = 8	Are grades comparable?
Written Q 1	M = 86.80 (SD = 7.01)	M = 79.13 (SD = 14.11)	t(1.8,42) = 1.47, **p = .179**
Written Q 2	M = 83.85 (SD = 12.56)	M = 85.94 (SD = 7.02)	t(1,26) = -.44, **p = .664**
Written Q 3	M = 83.30 (SD = 6.29)	M = 76.5 (SD = 14.26)	t(1,8.74) = 1.27, **p = .236**

As mentioned earlier, written quizzes were not the same tests so conclusions have to be tentative. The content of written quiz 3 is relatively similar in both conditions, however, since it is a final test that the instructor keeps and re-uses. It is about 75% similar with the exact same sections on discrimination and transcription. Results in Table 2 show no significant difference between the online class and the F2F class on any of the written quizzes.

Comparison of 7 oral quizzes

Table 3
Comparison of oral quizzes

	F2F Mean (SD)	Online Mean (SD) n = 8	Are grades comparable?
Oral Q 1	M = 83.85 (SD = 7.82), n = 17	M = 83.39 (SD = 5.66)	t(1,23) = .178, **p = .860**
Oral Q 2	M = 90.67 (SD = 4.58), n = 15	M = 86.38 (SD = 6.00)	t(1,21) = 1.92, **p = .068**
Oral Q 3 *	M = 88.94 (SD = 6.38), n = 19	M = 88.63 (SD = 3.85)	t(1,25) = .132, **p = .896**
Oral Q 4	M = 91.47 (SD = 4.39), n = 19	M = 86.00 (SD = 5.01)	t(1,25) = 2.84, **p = .009**
Oral Q 5*	M = 90.75 (SD = 5.00), n = 16	M = 84.43 (SD = 6.70)	t(1,22) = 2.60, **p = .016**
Oral Q 6*	M = 81.80 (SD = 78.25), n = 15	M = 78.25 (SD = 8.56)	t(1,21) = .686, **p = .500**
Oral Q 7	M = 88.01 (SD = 6.29), n = 17	M = 83.96 (SD = 6.52)	t(1,13.05) = 1.48, **p = .164**

Oral quizzes 3, 5, and 6, violated the assumption of normality as measured by the Shapiro-Wilk test: oral quiz 3 was not normal for the F2F class ($p = .027$); oral quiz 5 was not normal for the online class ($p = .017$); and oral quiz 6 was not normal for the F2F class ($p < .001$). For this reason, non-parametric tests were run for these three quizzes. Results are presented in Table 4.

Table 4
Non-parametric tests

Null Hypothesis	Test	Sig.	Decision
The distribution of OQ3 is the same across categories of Group.	Independent-Samples Mann-Whitney U Test	**.856**	Retain the null hypothesis.
The distribution of OQ5 is the same across categories of Group.	Independent-Samples Mann-Whitney U Test	**.027**	Reject the null hypothesis.
The distribution of OQ6 is the same across categories of Group.	Independent-Samples Mann-Whitney U Test	**.325**	Retain the null hypothesis.

As can be seen from Tables 3 and 4, only two oral quizzes showed a significant difference between both classes, namely oral quiz 4 ($p = .009$ in Table 3) and oral quiz 5 ($p = .027$ in Table 4).

Comparison of the evolution

To answer the second research question on the evolution of the pronunciation of the students and to gauge whether students in both conditions showed similar improvement, results on the first and last oral quizzes were compared. As mentioned above, only the students who turned in all seven recordings were included in this analysis. Results are presented in Table 5.

Table 5
Comparison of the evolution

	F2F Mean (SD), n = 17	Online Mean (SD), n = 8	Are grades comparable?
Oral Q 1	M = 83.85 (SD = 7.82)	M = 83.39 (SD = 5.66)	t(1,23) = .178, **p = .860**
Oral Q 7	M = 88.01 (SD = 6.29)	M = 83.96 (SD = 6.52)	t(1,13.05) = 1.48, **p = .164**
Difference	M = 4.16 (SD = 5.26)	M = .58 (SD = 3.68)	t(1,23) = 1.73, **p = .097**

The first two rows in this table are identical to those in Table 3. From there, the third row presents the difference between the two classes, with the conclusion that the difference is not significant (p = .097). In addition, to see the trajectory of students over the span of all seven assignments, the following figure presents the means for all seven oral quizzes given in Table 3. The two oral quizzes that were found to be significant are shown with an asterisk.

Figure 1
Trajectory of student grades

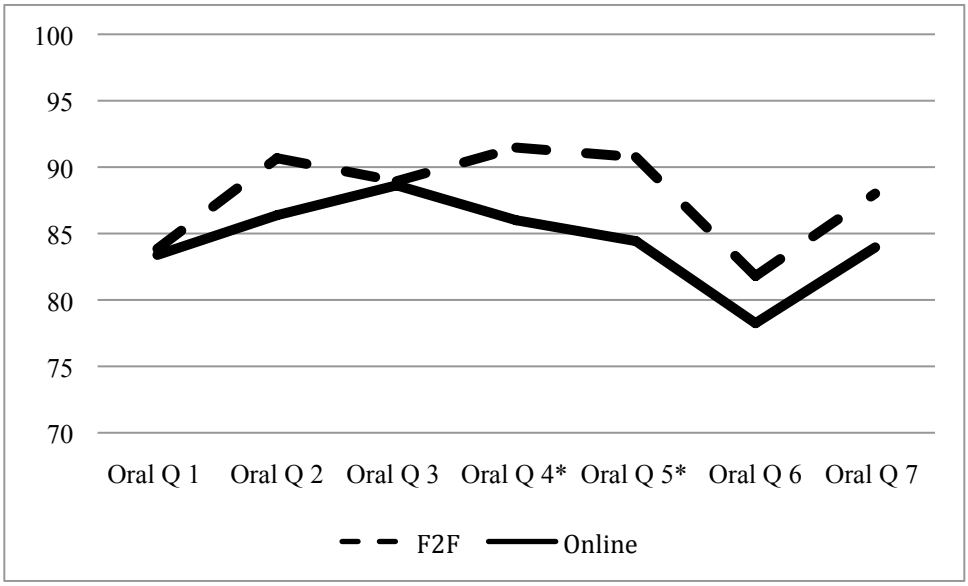


Figure 1 shows similar trajectories for both classes. The means for the online class are almost always lower, except for the first assignment, which helps us be confident that both groups of students were comparable at the beginning of the semester. As mentioned above, for oral quizzes 4 and 5, the online class received significantly lower grades. A discussion of the similar results for oral quiz 3 is presented in the following section.

DISCUSSION

As a reminder, the following research questions were asked:
1. Are the grades on assignments comparable in the F2F class and online class?
2. Is the evolution of the pronunciation of students between the first recording and last one comparable?
The first two sections below will provide an answer to the first research question and the third section to the second question.

Written quizzes

As seen in Table 2, there was no significant difference on written quizzes between the F2F and the online classes. However, since these were not identical tests (except for most

of the third test), conclusions have to be tentative. What we must cautiously recognize, however, is the possibility that learners performed in similar fashion on written tests regardless of the format of the class.

Oral quizzes

The exact same oral assignments were given to both classes so we can analyze the results provided in Tables 3 and 4 with more confidence. These two tables show a significant difference on two of the seven assignments given. Oral quiz 4 showed that the online class performed significantly lower than the F2F class. This quiz covered nasal vowels, which are perceptually and articulatorily difficult for students. They typically struggle to perceive the difference between [õ] and [ɑ̃], and to produce an accurate [ɑ̃], which they tend to realize as a front vowel.

Oral quiz 5 was probably the most difficult assignment during the entire semester as it included both schwas (which are both conceptually and articulatorily difficult) and liquids (articulatorily difficult). Schwas are quite complex in French phonetics as they are governed by rules that seems arbitrary and unpredictable to students. In addition, the type of input that they were exposed to through teacher talk tends to produce more schwas than natural speech. As the class in general, and this assignment in particular, stressed a less formal and more natural style of speech, students tended to have trouble shedding habits of pronouncing schwas they have heard for years. Some students over my years of teaching of teaching this class have told me that they know the rules and understand them, but they just cannot stop themselves from saying the schwas they should not pronounce. In addition to this, oral quiz 5 included liquids. It is a well-known fact that most English-speaking students have trouble with the French /r/ sound. Having to focus on two complex items in one quiz may help explain why students in the online class performed at a significantly lower level than those in the F2F class.

It seemed that this difference may be due to the type of feedback that each class received. Sagarra and Zapata (2008, p. 219) reported that "when asked what they liked the most about the online workbook, participants responded that they enjoyed (…) receiving individualized immediate feedback" among other things. This was possible in the online format (but not for the F2F) for written activities, but impossible for oral ones since feedback, though personalized, was delayed. For F2F oral activities, feedback was immediate but not individualized so as to be less face-threatening. Sheen (2010, p. 225) claimed that the timing was not crucial and concluded that

> the crucial factor that influences the effectiveness of C[orrective] F[eedback] is the explicitness of the feedback (i.e., whether its corrective force is clear). (…) What is crucial is not whether the CF is on-line or offline but whether it is explicit. In other words, the current study suggests that it may not matter whether the CF is provided immediately or is delayed as long as it is noticeable.

Since both formats received feedback, the immediate feedback that students in the F2F class seem to enhance accuracy in production, at least for the more complex themes. By

contrast, personalized feedback, as provided to the online class, may not be as effective because of the delay in receiving the feedback. By receiving immediate and explicit (positive or negative) feedback, students may be more able to notice the difference between their production and the model, hence making the corrections more salient and more meaningful, and ultimately more useful, than those receiving delayed feedback. This seems to contradict Metcalfe, Kornell and Finn (2009) for whom delayed feedback produced better performance than did immediate feedback but their study focused on a radically different skill of vocabulary learning.

Evolution through the semester

As we have seen , there was no statistical difference between the two classes in the overall improvement from oral quiz 1 to oral quiz 7. In addition, both classes show similar trajectories in their evolution of French pronunciation over the course of the semester. In Figure 1, we notice one exception to otherwise fairly parallel lines: for oral quiz 3, the F2F line shows a small dip that we do not see in the online class. This dip reflects means scores that are almost identical: 88.94 for the F2F class vs. 88.63 for the online class. The phonetic theme covered in this oral quiz is liaison, a theme that I deem the most complex in the semester and to which three entire classes are devoted. By comparison, schwas are studied over the course of two class periods. As we have seen, the oral quiz with schwas (and liquids) shows a significantly higher performance by the F2F class, a result which was explained in part by the difficulty of the theme and the type of feedback. It, therefore, seems all the more surprising that we do not see the same pattern with liaison. In fact, it could be the case that, in the case of liaison, immediate feedback (most of which is negative feedback followed by a repetition of the rule) is counter-productive to learning. Indeed, in class, students often seem overwhelmed by their perceived inability to achieve a correct result and look like they tune out. By contrast, students in the online class have control over their learning, reviewing rules when they feel they need it, retrieving feedback on their own time rather than being forced to get it from the instructor in class. This may allow them more time to process the material as well as make a deeper connection with the material.

CONCLUSION AND QUALITATIVE PERSPECTIVES

Overall, the comparison between the F2F and the online class shows rather positive results. In other words, the online class is not significantly lower than the F2F class on all but two assignments. It even seems to have an advantage over the F2F class for one oral assignment. These results, however, can be disappointing for the instructor when the instructor's time commitment is factored in. Not only did it take a very large amount of time to develop this online class, but the amount of time spent on grading is also much higher than in the F2F class, partly because of the increased difficulty of working with IPA symbols on a keyboard and partly because of the focus on individualized feedback. Students often seem to have the perception that an online class is easier and less-work intensive than a F2F class. Some administrators and instructors may also think that, since everything is computer graded, larger class sizes are warranted. While these might be true in some cases, it was not the experience of this instructor. Although activities can be re-

used for future classes (with minor changes at most), the amount of time for grading is not likely to be radically different, if the instructor wants to keep the integrity of the student experience. Indeed, when the online class was taught again during the summer of 2014, the maximum enrollment was set at 20 (instead of 26 for the F2F class).

In spite of the focus on individualized feedback and the fairly positive results, students enrolled in the online class seemed to think that the online format was not as good as a F2F class. The following three quotes come from a questionnaire given mid-semester and at the end of the semester to online students. They provide insights on the perceptions of the class and reflect one of the problems that other researchers have observed, namely the lack of connection and social community. They are another reason why the results of this study were deemed to be 'rather positive' instead of 'positive'.

> "I feel like my speaking didn't develop as well as it could have. I skyped with the professor when I had questions and that was helpful but I don't think my speaking improved very much from chapter to chapter. If I was practicing more regularly and hearing the professor and other students speak in a classroom setting I think I would have absorbed more."
>
> "I missed having fellow students to work with and just commiserate. There's something about being able to roll one's eyes together when you're reminded about a quiz or asking each other what you got on a test that feels vital to me. I guess I just missed the social dimension of the classroom."
>
> "I feel that it would have been very beneficial to me to have seen or heard other students and their progress as well."

These comments point to the importance of developing a sense of community and connectivity as discussed earlier. Group work was eliminated from the online class compared to the F2F class, but this may have been a mistake. Instead, requiring students to use tools such as Skype, Facebook or Google Hangouts to connect with each other on guided activities may be beneficial.

ACKNOWLEDGEMENTS

The author wishes to thank Le Anne Spino (from Second Language Studies program at Michigan State University) for her help with statistics, and CeLTA (Center for Language Teaching Advancement) at Michigan State University for the Language Learning in Online Environments grant that made this study possible.

ABOUT THE AUTHOR

Anne Violin-Wigent is an Associate Professor of French at Michigan State University. She receives her Ph.D. in French from Purdue University in 2001. She teaches French phonetics and pronunciation, introduction to French linguistics, and graduate seminars on

the history of the French language, spoken French, and gastronomy. Her research focuses on sociophonetics and leveling in France. She recently co-authored a textbook on French pronunciation *Sons et Sens: La prononciation du français en contexte (*2013) with Jessica Miller and Frédérique Grim.

REFERENCES

Abuseileek, A. F. (2009). The effect of using an online-based course on the learning of grammar inductively and deductively. *ReCALL, 21*(3), 319-336.

Blake, R. (2008). *Brave new digital classroom: Technology and foreign language learning*. Washington, D.C.: Georgetown University Press.

Blake, R. (2009). From web pages to distance learning: Technology in the foreign language curriculum. In Lancashire, I. (ed). (2009). *Teaching Literature and Language Online*. New York: Modern Language Association of America. 23-37.

Blin, F. ,& Appel, C. (2011). Computer supported collaborative writing in practice: An activity theoretical study. *CALICO Journal, 28*(2), 473-497.

Brudermann, C. (2010). Promoting EFL phonological skills in a blended learning environment: Pedagogical considerations for online tutoring practices amongst non specialist students. *Les Cahiers de l'Apliut, 29*(3), 50-65.

Carnevale, Dan. 2007. Distance learning going the distance. http://chronicle.com/blogs/wiredcampus/distance-learning-going-the-distance/3422

Eneau, J., & Develotte, C. (2012). Working together online to enhance learner autonomy: Analysis of learners' perceptions of their online learning experience. *ReCALL, 24*(1), 3-19.

Goertler, S. & Winke, P. (ed). (2008a). *Opening doors through distance language education: Principles, perspectives, and practices*. San Marcos: CALICO.

Goertler, S., & Winke, P. (2008b). The effectiveness of technology-enhanced foreign language teaching. In Goertler, S. & Winke, P. (ed). (2008a). *Opening Doors through Distance Language Education: Principles, Perspectives, and Practices* (pp. 233-259). San Marcos: CALICO.

Lancashire, I. (ed). (2009). *Teaching literature and language online*. New York: Modern Language Association of America.

McBride, K., & Fägersten K. B. (2008). Student's role in distance learning. In Goertler, S. & Winke, P. (ed). (2008a). *Opening Doors through Distance Language Education: Principles, Perspectives, and Practices,* (pp. 233-259). San Marcos: CALICO.

Metcalfe, J., Kornell, N., & Finn, B. (2009). Delayed versus immediate feedback in children's and adults' vocabulary learning. *Memory & Cognition*, *37 (8)*, 1077-1087.

Oliver, K., Kellog, S., & Patel, R. (2012). An investigation into reported differences between online foreign language instruction and other subject areas in a virtual school. *CALICO Journal, 29*(2), 269-296.

Russell, V., & Curtis, W. (2013). Comparing a Large- and small-scale online language course: An examination of teacher and learner perceptions. *The Internet and Higher Education, 16*, 1-13.

Sagarra, N., & Zapata, G. (2008) Blending Classroom instruction with online homework: A study of student perceptions of computer-assisted L2 learning. *ReCALL* 20.2: 208-224.

Sanders, R. F. (2005). Redesigning introductory Spanish: Increased enrollment, online management, cost reduction, and effects on student learning. *Foreign Language Annals* 38.4: 523-532.

Senior, R. (2010). Connectivity: A framework for understanding effective language teaching in face-to-face and online learning communities. *RELC Journal* 41.2: 137-147.

Sheen, Y. 2010. Differential effects of oral and written corrective feedback in the ESL classroom. *Studies in Second Language Acquisition* 32.2: 203-234.

Violin-Wigent, A. Miller, J. & Grim, F. (2013). *Sons et Sens: La prononciation du français en contexte.* Washington D.C.: Georgetown University Press.

White, C. (2003). *Language learning in distance education*. Cambridge: Cambridge University Press.

Wildner-Bassett, M. (2008). Teacher's role in computer-mediated communication and distance learning. In Goertler, S. & Winke, P. (ed). (2008a). *Opening Doors through Distance Language Education: Principles, Perspectives, and Practices* (pp. 67-83). San Marcos: CALICO.

Steed, W., & Delicado Cantero, M. (2014). Fair Dinkum: L2 Spanish pronunciation in Australia by the book. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 58-67). Ames, IA: Iowa State University.

FAIR DINKUM: L2 SPANISH PRONUNCIATION IN AUSTRALIA BY THE BOOK

William Steed[1] and Manuel Delicado Cantero[2]
James Cook University[1] and Australian National University[2]

While Spanish teaching in the United States is a well-established field, it is still maturing in Australia. Clear evidence of this is the absence of materials aimed at the Australian English-speaking market. While for the most part foreign textbooks are fine, problems arise immediately when it comes to addressing L2 Spanish pronunciation in a pronunciation course, as the models and examples are aimed at an audience with a different language or dialect background. The result is that students report it as difficult to get the most benefit from the materials. Just like the pedagogy of general phonetics is best achieved using Australian-focused materials, the same focus is necessary for effective L2 pronunciation materials. This study aims to review well-known textbooks for pronunciation courses, all primarily and even explicitly written for American students – arguably the largest market for Spanish pedagogical materials. Our review only measures appropriateness to the Australian university classroom context not the quality of the texts. We divide our approach into two main aspects: phonological and phonetic content (segmental and suprasegmental) and level of appropriateness to the linguistic awareness of Australian learners of Spanish.

INTRODUCTION

The US market for learning Spanish is much greater than that of Australia. Its proximity to Spanish-speaking countries and its own Spanish-speaking population mean that most residents are exposed to Spanish frequently through their lives. Unsurprisingly, it has become a leading producer of materials for learning Spanish. Although Spanish is one of the most widely spoken non-English languages in Australia (Australian Bureau of Statistics, 2013), there are comparatively fewer Australian learners of Spanish (Ministerio de Educación, Cultura y Deporte, Spanish Government 2012), and, as a result, few resources specifically targeted to them.

Grammatical instruction materials targeted at North American, English-speaking learners are generally suitable for Australian learners of Spanish – there are no major syntactic or morphological differences between US and Australian English standards. Pronunciation, on the other hand, differs substantially. Although US and Australian English are easily mutually intelligible, the differences lead to different accents in Spanish and different needs when working towards an acceptable and comprehensible pronunciation. Thus, materials produced for American learners of Spanish pronunciation will not necessarily target the phonetic and phonological characteristics of an Australian-Spanish interlanguage.

Furthermore, North American and Australian university majors are also structured differently, especially in the number of subjects per major, potentially leading to relevant differences in knowledge of the target language.

In light of the absence of Australian-targeted pronunciation pedagogy resources, the aim here is working towards meeting the needs of the Australian students: to review the suitability of

Australian learners of Spanish. Our research question is, to what extent are US-targeted Spanish pronunciation textbooks suitable for use in Australian university classrooms? We must emphasise that it is not the intent here to review the pedagogical styles, nor the quality of the content per se, but only to view them in the context of an Australian Spanish university classroom in order to locate and discuss any potential difficulties.

METHODOLOGY

Selection

The texts in question were chosen by several criteria:

- intended use as a pronunciation textbook for a pronunciation/phonetics course
- availability
- popularity
- publication date
- publication place
- written in Spanish

It is practical to review texts that will be easily accessible in a Spanish university classroom in Australia, thus availability in the Australian market is an important factor. A recent publication date (2003 or later) ensures that only current and up-to-date publications are considered. The publication place criterion considers only textbooks written for the large North American market, rather than those marketed to Spanish learners who are not necessarily English speakers. Choosing popular texts (that is, those that are currently used in North American Spanish programs) ensures that we are comparing texts that are considered appropriate in their target market. We selected only books written in Spanish as it is expected that the class will be taught in the target language. Other texts may be equally or more appropriate choices as pronunciation textbooks, but do not contribute to answering the question at hand. The five books reviewed in this article matching the criteria are as follows:

Guitart, J. M. (2004). Sonido y Sentido: Teoría y práctica de la pronunciación del español. Washington DC: Georgetown University Press.

Morgan, T. (2010). Sonidos en contexto: Una introducción a la fonética del español con especial referencia a la vida real. New Haven, CT: Yale University Press.

Piñeros, C. (2008). Estructura de los sonidos del español. Upper Saddle River, NJ: Prentice Hall.

Schwegler, A., Kempff, J., and Ameal Guerra, A. (2010). Fonética y fonología españolas, Hoboken, N.J.: Wiley.

Stokes, J. D. (2005). ¡Qué bien suena!: Mastering Spanish Phonetics and Phonology. Boston, MA: Houghton Mifflin.

Analysis criteria

The texts were rated according to two sets of criteria: content and context appropriateness. The content criteria are designed to determine whether the texts cover the topics that Australian learners of Spanish need to know in order to reach an appropriate level of Spanish pronunciation. The context appropriateness criteria examines them within a typical Australian

The content criteria were developed through auditory and basic acoustic analysis of typical Australian L2 Spanish learners, collected from recordings of end-of-semester oral examinations of first- and second-year learners of Spanish at an Australian university, and self-recordings of second- and third-years students of a single text as baseline benchmarks for a pronunciation course. Recurring errors were regarded as necessary features of pronunciation instruction. Cox (2012) served as the reference for Australian English pronunciation. We also considered attitudinal survey data from students at the same university with regards to the difficulty of aspects of pronunciation, triangulating it with the recorded data, and anecdotal reported errors from teachers of Spanish working in Australia.

The context criteria are collated from several sources. Survey data provided information about the linguistic experience of Spanish students at an Australian university (see Steed & Delicado Cantero, 2014). This was supplemented by an informal investigation of Spanish language major requirements at nine Australian university websites. The criteria used here examine the prerequisite knowledge of the textbooks in terms of metalinguistic awareness, conscious linguistic knowledge, level of Spanish and cultural knowledge for both Spanish speaking countries and the United States. It also examines the diversity of examples given, in light of the different cultural make-up of Spanish speakers that students will encounter in Australia compared to the US.

RESULTS

Segmentals

The simple articulation of vowels is an essential part of understanding Spanish pronunciation. Although we expected that all texts would include basic information about the articulation of Spanish vowels, we include it to ensure that the texts cover the simplest aspects for the student. The three most common vowel differences between L1 Spanish and Australian L2 Spanish are neutralisation, diphthongisation of monophthongs, and different realisation of diphthongs. All three are also common to US L2 Spanish, but the realisations of diphthongs differ in the two L2 interlanguages.

Table 1 shows the results for vowel content in each text. Each text has detailed description of the articulation of vowels and diphthongs in Spanish. The texts all describe the relative height, backness and rounding of the vowel phonemes. Where examples are given, they are typically US English, e.g. Schwegler et al.'s (2010) comparing Spanish /e/ to the vowel in 'bait' (in US English typically [e ~ eɪ], but [æɪ] in Australian English), Spanish /u/ to the vowel in Sue ([u ~ ʉ] in the US and [ʉː] in Australia), and Spanish /o/ to the vowel in 'sow' (in US English typically [o ~ oʊ], but [əʉ] in Australian English) (see Cox 2012).

Each text also advises readers about two common vowel errors in US and Australian L2 Spanish – neutralisation of vowels (particularly /e/) in unstressed syllables and diphthongization of mid-vowels at the end of words. However, none of the texts has exercises for the realisation /eɾ/ as /ɜɾ/ in stressed syllables typical of Australian English.

As for consonants, we selected seven features for evaluation, primarily about phonetic realisations of phonemes rather than phonological contrasts. Although each of the texts includes descriptions of each phoneme, most of the L1 Spanish/Australian L2 Spanish differences do not relate to neutralisation of phonological contrasts.

Table 1
Information on segmentals

Textbook	*VOT*	*Place of articulation*	*Rhotic articulation*	*Approximant [β ð ɣ]*	*Post-vocalic /r/*	*/t d/ neutralization*	*/s/ aspiration*	*Articulatory differences (vowels)*	*Neutralization (schwa)*	*Diphthongization*	*Diphthong comparison (Eng – Span)*
Guitart	2	2	2	2	**0**	2	2	2	2	2	**1**
Morgan	2	2	2	2	**0**	2	2	2	2	2	**0**
Piñeros	2	2	2	2	**0**	2	2	2	2	2	**0**
Schwegler et al.	2	2	2	2	**0**	2	2	2	2	2	2
Stokes	2	2	2	2	**0**	2	2	2	2	2	2

Note. 0=content absent; 1=some information or indirect information; 2=clear information and explanation (with varying degrees of depth).

All five texts have excellent coverage of most aspects of consonant articulation relevant to Australian Spanish learners, with one important exception: the lack of information and practice for /ɾ/-deletion in the syllable coda. While this is understandably the result of most North American English dialects being rhotic, it is characteristic of Australian L2 Spanish.

Suprasegmentals

The suprasegmental content (see Table 2) results contrast sharply with recent research, which remarks the need to pay due attention to prosodic information in the language classroom at all levels (Cortés, 2002; Lahoz, 2012; Santamaría, 2007), including the recommendation that prosody be the first content to be covered and practiced in class (Gil, 2007; Lahoz, 2012; see also the Curriculum Plan for the Instituto Cervantes in Instituto Cervantes, 2007). The content and exercises contained in Lahoz (2012) and the series of online materials in the Centro Virtual Cervantes – e.g., Santamaría (2008a, b) and their continuations – serve as models for the teaching of suprasegmentals in a Spanish class.

Table 2
Information on suprasegmentals

Textbook	*Intonation*	*Neutral intonation*	*Emotive intonation*	*Sentence level stress*	*Lexical stress*	*Tonic group/phonological word*	*Phonic group, pause*	*Rhythm*	*Syllabification*	*Synalepha, synaeresis*	*Resyllabification*
Guitart	2	**1**	**0**	**0**	2	2	2	**1**	2	2	2
Morgan	**0**	**0**	**0**	**0**	2	**1**	**0**	**0**	2	2	2
Piñeros	2	2	2	**0**	2	2	2	**0**	2	2	2
Schwegler et al.	2	2	**1**	**0**	2	**1**	2	2	2	2	2
Stokes	2	2	**0**	**0**	2	**1**	**0**	2	2	2	2

Note. 0=content absent; 1=some information or indirect information; 2=clear information and explanation (with varying degrees of depth).

We note variable coverage of prosodic information and practice. On the one hand, all cover syllabification and resyllabification, synalepha/synaeresis, and lexical stress; the latter is usually linked to accent mark explanations, as is the case in Morgan (2010, chapter 4), for instance. On the other hand, however, most lack adequate information on pauses, phonic groups (that is, prosodic intonation groups), tonic groups and the phonological word, and sentence level stress, content that is necessary for any Spanish student across the world. In certain cases, the information is introduced briefly: tonic groups in Stokes (2005, pp. 52-53) are introduced indirectly by way of listing categories that will usually be unaccented (clitic pronouns, short prepositions, determiners, etc.). While almost all of them cover intonation, only two go beyond the basic neutral use of intonation (declarative sentence, basic interrogation, etc.) to cover emotive uses. Piñeros (2008, chapter 23) abounds in details, while Schwegler et al. (2010, pp. 325-326) is brief. Finally, information and practice on rhythm and tempo are present in three textbooks, again with varying degrees of detail.

While this problem will be equally important for US learners, Australian learners have less exposure to Spanish in their everyday life and thus potentially lower awareness of the prosodic structure of L1 Spanish. This may require more specific instruction in the classroom.

Context

For a pronunciation course to be successful, it is important to discern whether students will be able to a) understand the Spanish level of the text, and b) follow the linguistic terminology used in the text. It is also important to the students that the content is covered to an appropriate depth for an Australian university level course.

In addition, the level of Americocentrism was gauged by considering to what extent the examples occurring in the texts used North American Spanish, and whether the examples had the potential to be confusing for a non-linguistically aware Australian English speaker.

Some of these criteria are subjective in measure and are judged from the researchers' experience in Spanish learners' reading and comprehension ability at the prerequisite level for a pronunciation course. The results are summarized in Table 3:

Table 3
Information on context

Textbook	*Variable English in examples*	*Explanation of phonology*	*Explanation of phonetics*	*Articulation*	*Spanish dialectal variation*	*Diversity of Spanish examples*	*Glossary of terms*	*No linguistics background*	*Spanish proficiency*	*No specific English as model*	*No specific student target*
Guitart	**0**	2	2	2	2	2	2	**0**	2	**0**	**0**
Morgan	**0**	2	2	2	2	2	2	**1**	**1**	**0**	**0**
Piñeros	**0**	2	2	2	2	2	**0**	**0**	2	**1**	**0**
Schwegler et al.	**0**	2	2	2	2	2	2	**1**	2	**0**	**0**
Stokes	**0**	**1**	2	2	2	2	2	**1**	**1**	**0**	**0**

Note. 0=content absent; 1=some information or indirect information; 2=clear information and explanation (with varying degrees of depth).

With regards to level of Spanish, Spanish university majors in Australia are not completely comparable to US college majors. Firstly, typical Australian university majors comprise 6-9 semesters of study (300-450 hours) across three to four years. Even as an advanced subject, Australian students taking a specific pronunciation course may have as little as three semesters (150 hours) of instruction experience in Spanish.

Secondly, while in the US basic language classes are typically part of the general education courses (GECs) before declaring and thus beginning the actual major, no such requirement exists in Australian degrees. A practical difference is that while a US student may not be allowed to count any credits taken as GECs as part of their Spanish major, an Australian student will typically include all university-level credits towards their major. Our own review of Spanish programs in Australia shows that a typical language major consists of about 7 or 8 subjects, including basic language classes, much less than a typical US major.

Thirdly, in the case of the US, while there are there are too many programs to perform an exhaustive analysis, all programs will include at least some type of advanced grammar as part of their requirements. An online look at a number of well-known programs in Spanish in some US universities (Ohio State, University of New Mexico, UCLA, University of Wisconsin-Madison, Pennsylvania State, Indiana, University of Illinois at Urbana-Champaign, and University of Texas-Austin) reveals that it is clear that a Spanish pronunciation course of some denomination is present, at least as an elective, if not as a requirement for all students declaring a Spanish major or minor. The existence of these courses explains the existence of the textbooks. While we are aware that not all US Spanish programs require or even offer a pronunciation course, it is clear that this is an offering that is common and expected. Such is not the case in Australia, given the competence-focussed nature of Australian language majors.

Fourthly, as already mentioned, Australian learners of Spanish have a much lower level of exposure to L1 Spanish compared to their North American counterparts. While there is a substantial Spanish-speaking community in Australia (ABS, 2013), there is far less widespread public use of Spanish, for example in media and social interaction.

As a result of these differences, the instructor needs to take into consideration that the level of Spanish of the average Australian student is likely to be lower than that of US learners. A tangible consequence of this is that the level of Spanish in three of the texts, namely Piñeros (2008), Guitart (2004) and possibly Schwegler et al. (2010), will be difficult for students, especially those who are struggling. The level of Spanish in Stokes (2005) is at a more appropriate level for the students who are enrolled in Australian pronunciation courses.

Another conflictive fact has to do with the English used in the textbooks. As they are targeted at American students, it comes as no surprise that the model is US English. However, this fact leads to difficulties when it comes to benefitting from the input and especially from the comparisons English-Spanish a number of exercises rely on. When vowels, alveolar trill in coda, dark /l/, etc. are put together, the result is the decreased effectiveness of a number of sections and exercises in the books. A case in point from our teaching experience is the English/Spanish contrast in Morgan (2010, p. 346); students were not satisfied with the American English models for comparison.

The instructor of a pronunciation course will also find that linguistics is not a required field of study for Spanish students in most Australian Spanish majors. We surveyed Spanish students at one university, finding that only 1/3 of students at all levels were taking or had taken a course in linguistics (Steed & Delicado Cantero 2014). It cannot be expected, then, that students have metalinguistic awareness, particularly surrounding pronunciation. While they may have learned about some topics (such as English syllables, stressed syllables, rhyming, etc.) in school, students are not typically accustomed to critically assessing pronunciation,

and most linguistic concepts (phonetic transcription, phonemes, intonational awareness, etc.) must also be explicitly taught.

The texts, for the most part, explain the vocabulary and concepts that students typically find unfamiliar, in varying detail: while Piñeros (2008) is thorough and complex, Stokes's (2005) explanations are briefer. Save for Piñeros (2008), all include a glossary.

Australian universities are regulated by the Tertiary Education Quality Standards Assurance (TEQSA) using the Australian Qualifications Framework (AQF). Undergraduate courses are expected to lead to level 7 outcomes, "broad and coherent knowledge and skills for professional work and/or further learning" (AQF, 2013). It is thus important that a text used to teach pronunciation to higher level students be aiming for level 7 comprehension of the topic.

CONCLUSIONS

As stated earlier, the aim of this study is not to review the textbooks for pedagogical quality, nor is it to assess how well the texts achieve their own aim – that is, teaching North American students about Spanish phonetics and phonology. Our aim is to assess how well the textbooks can be adapted to a purpose they were not specifically designed for – teaching Australian students.

Our results show that the texts do not cover all of the typical errors that Australian students make in pronouncing Spanish. In addition, examples given from US English may mislead Australian learners without expert guidance from their instructor. Learners may even unintentionally acquire mismatched correspondences between English and standard Spanish pronunciation and, in our experience, will also find the US English input inadequate to their purpose. Furthermore, without the background of basic linguistics, which is not typical for students of Spanish in Australia, or the higher level of Spanish that can be expected toward the end of a US Spanish major, the texts may be too difficult to follow, or have too much content for a single course's content. On the other hand, a simpler textbook may not have a deep enough content to satisfy the higher cognitive requirements of an Australian university-level course. The result is a conflict between appropriate level of Spanish and presumed knowledge of linguistics and depth and breadth of explanations and activities. We can confirm, then, that there is a substantial gap which needs to be covered in order to use these texts effectively in Australian subjects.

Teachers who are choosing an appropriate text can use the results of our investigation to guide their selection. Our conclusions suggest that, although the texts are of good quality and teachers can use any of them in a pronunciation class, no one text examined stands out as particularly better than the others in terms of content and context appropriateness. A student needs analysis for their specific teaching context is best to guide them to the best choice for their students. Due to the lack of Australian-focussed materials, however, Australian teachers will need to supplement their choice with additional materials.

ABOUT THE AUTHORS

William Steed is Lecturer in Linguistics for the Speech Pathology program at James Cook University. He received his PhD from the Australian National University in 2011 for his

teaching and research, along with the fact of being both a native speaker of Australian English and a speaker of L2 Spanish, lead him to co-found the *Australian Spanish Pronunciation Pedagogy Project* to improve the state of research into L2 Spanish in Australia.

Manuel Delicado Cantero is Lecturer of Spanish at the Australian National University. As part of his research-led teaching, he co-founded the *Australian Spanish Pronunciation Project*, which aims to research L2 Spanish pronunciation in this part of the world. The researchers in this team seek to understand the production and perception issues Australian students face in this area and to produce the necessary materials to meet their needs. His research interests also include Spanish formal syntax and historical syntax.

REFERENCES

Australian Bureau of Statistics. (2013). *Cultural diversity in Australia: Reflecting a nation, stories from the 2011 census*. Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2071.0main+features902012-2013

Australian qualifications framework. (2013). *AQF Levels*. Retrieved from http://www.aqf.edu.au/aqf/in-detail/aqf-levels

Cortés, M. (2002). *Didáctica de la prosodia del español: la acentuación y la entonación*. Madrid: Edinumen.

Cox, F. (2012). *Australian English. Pronunciation and transcription*. New York: Cambridge University Press.

Gil Fernández, J. (2007). *Fonética para profesores de español: De la teoría a la práctica*. Madrid: Arco/Libros.

Guitart, J. M. (2004). *Sonido y Sentido: Teoría y práctica de la pronunciación del español*. Washington DC: Georgetown University Press.

Instituto Cervantes (2007). *Plan curricular del Instituto Cervantes. Niveles de referencia para el español.* Retrieved from http://cvc.cervantes.es/Ensenanza/Biblioteca_Ele/plan_curricular/default.htm

Lahoz, J. M. (2012). La enseñanza de la entonación, el ritmo y el tempo. In J. Gil (Ed.), *Aproximación a la enseñanza de la pronunciación en el aula de español* (pp. 93-132). Madrid: Edinumen.

Ministerio de Educación, Cultura y Deporte (Government of Spain) (2012). *El mundo estudia español 2012*. Retrieved from http://www.mecd.gob.es/redele/el-mundo-estudia-espa-ol/2012.html

Morgan, T. (2010). *Sonidos en contexto: Una introducción a la fonética del español con especial referencia a la vida real*. New Haven, CT: Yale University Press.

Piñeros, C.-E. (2008). *Estructura de los sonidos del español*. Upper Saddle River, NJ: Prentice Hall.

Santamaría Busto, E. (2007). Enseñar prosodia en el aula: reflexiones y propuestas. In E. Balmaseda Maestu (Ed.), *Las destrezas orales en la enseñanza del español L2-LE: XVII Congreso Internacional de la Asociación del Español como lengua extranjera (ASELE)* (Vol. 2, pp. 1237-1250). Retrieved from http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/17/17_1237.pdf

Santamaría Busto, E. (2008a). *La palabra fónica y el ritmo (I)*. Retrieved from http://cvc.cervantes.es/aula/didactired/anteriores/mayo_08/26052008a.htm

Santamaría Busto, E. (2008b). *Un café con entonación... perdón... con leche*. *La entonación I*. Retrieved from http://cvc.cervantes.es/aula/didactired/anteriores/julio_08/0707 2008a.htm

Schwegler, A., Kempff, J., & Ameal Guerra, A. (2010). *Fonética y fonología españolas*, Hoboken, N.J.: Wiley.

Stokes, J. D. (2005). *¡Qué bien suena!: Mastering Spanish phonetics and phonology*. Boston, MA: Houghton Mifflin.

Steed, W., & Delicado Cantero, M. (2014). *First things first: Exploring Spanish students' attitudes towards learning pronunciation in Australia*. Manuscript submitted for publication.

Zetterholm, E., & Tronnier, M. (2014). Different stress patterns meet: Kurdish L1 speakers learn Swedish. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 68-74). Ames, IA: Iowa State University.

DIFFERENT STRESS PATTERNS MEET: KURDISH L1 SPEAKERS LEARN SWEDISH

Elisabeth Zetterholm, Linnaeus University
Mechtild Tronnier, Lund University

The Kurdish language is an Indo-European language spoken in different countries from a region in the Middle East. Many speakers of Kurdish have migrated to other countries, including Sweden, and it is a frequent L1 spoken in Sweden. More than 3000 students with Kurdish L1 were registered as adult learners of Swedish for immigrants in the year 2012 (www.scb.se). For L2 learners it is important to be aware of differences in phonology and prosody between L1 and L2 to reduce problems in direct communication. In Swedish, a stressed syllable is perceived through intensity, pitch and duration. The stress placement can change depending on the morphology. In compounds, the main stress is on the first element and a secondary stress on the last element. In Kurdish, on the other hand, stress is not correlated with length variation and stress is usually on the last syllable. Compounds only have one stressed syllable on the last element. The investigation of Kurdish L1 speakers learning Swedish as an L2 indicates potential problems with stress placements, which causes problems in communication since the stressed syllable serves as a perceptual anchor for a native listener.

INTRODUCTION

Kurdish is spoken by approximately 20,000,000 speakers in a region which is part of five different countries, namely Turkey, Iran, Iraq, Syria and Armenia. Kurdish is a minority language in these countries and is divided into different varieties with the two major dialects Sorani and Kurmanji (Hassanpour, 2012). The dialect Sorani is one of the official languages in Iraq (Hassanpour, 2012). The alphabet of the Sorani dialect, spoken mainly in the southern part of Kurdistan, is based on the Arabic alphabet, while the alphabet of the Kurmanji dialect, spoken mainly in the northern part, is based on the Latin alphabet (Hassanpour, 2012). Speakers of one dialect usually understand speakers of other Kurdish dialects. Kurdish belongs to the Iranian branch of the Indo-European language family and is related to dialects spoken in the surrounding areas.

When it comes to learners of Swedish as a second language in Sweden, Kurdish is among the ten most frequent L1s. Thus, it is of interest to know if there are specific problems for Kurdish speakers related to their native language. If so, that should be highlighted in the teaching situation for Kurdish L1-speakers. This study is part of a research project about how learners with different native languages acquire and learn the pronunciation in Swedish as a second language (e.g. Tronnier & Zetterholm, 2011). This is important since the phonology and the prosody of L1 often seem to have an impact on how easy or difficult it is to acquire an intelligible pronunciation in an L2 (e.g. Abrahamsson, 2009; Bannert, 2004; Munro, 2008).

A comparison of stress patterns

Stress patterns in Swedish

A distinctive prosodic feature of Swedish prosody is word stress, which can occur on any syllable depending on the morphology. According to Bruce (2012), there are two basic word stress patterns in Swedish. The *simple* stress pattern is characterized by one primary stress within any of the last three syllables in a word. The *complex* stress pattern is characterized by one primary stress early in the word and a secondary stress later in the word, usually on the last part. Examples from standard Swedish on the distribution of primary and secondary stress in compounds are: ˈ*kaffeˌbryggare* (coffee machine) and *moˈbilteleˌfon* (cell phone) and ˈ*födelsedagskaˌlas* birthday party. Varieties of stress placement can be found in different regional dialects, especially in the northern part of Sweden. In Swedish there are some minimal pairs with distinctions in meanings depending on the stress, e.g., *formel* [ˈfɔrmɛl] (formula) – *formell* [fɔrˈmɛl] (formal); *kaffe* [ˈkafə] (coffee) – *café* [kaˈfeː] (café); *armen* [ˈarmɛn] (the arm) – *armén* [arˈmeːn] (the army).

The syllable structure of Swedish permits as many as three initial consonants and three final consonants in root morphemes together with one vowel (Garlén, 1988). One long single vowel can also occur as one syllable e.g., *i* [iː] (in), *ö* [øː] (island), *å* [oː] (small river). In Swedish there is a distinction concerning duration in syllables (Bruce, 2012). There is a complementary quantity distinction in stressed syllables, expressed by variation of segmental length namely (C)V:(C) or (C)VC:(C). The stressed syllables are also produced with a slightly higher intensity and pitch compared to other syllables in the word or phrase. Variation in vowel quantity also brings along a difference in vowel quality (Bruce, 2012; Riad, 2013), e.g. *bus* [bʉːs] (mischief) – *buss* [bʊsː] (bus); *hat* [hɑːt] (hate) – *hatt* [hatː] (hat); *mål* [moːl] (goal) – *moll* [mɔlː] (minor key); *stöta* [støːta] (bump) – *stötta* [stœtːa] (support).

Stress patterns in Kurdish

The syllable structure allows two initial consonants and three final consonants, (C)CV(C)(C)(C) in Kurdish. A syllable has at least one consonant and one vowel, but one single vowel cannot be in isolation (Rahimpour & Dovaise, 2011). In Kurdish, stress is not correlated with length or quality variation in the syllable, but loudness and pitch has a strong effect (Rahimpour & Dovaise, 2011). Stress in Kurdish is often predictable. Only one syllable is stressed in a multisyllabic word, whereas the others are unstressed. Mostly, the stress falls on the final syllable and in nominal compounds the stress is on the last element. However, for some words the stress can change depending on the morphology, and some morphemes take stress, e.g. ˈ*hatin* (they came) – *haˈtin* (to come) (Rahimpour & Dovaise, 2011).

The study

Only read speech is used in this study. The participants read prepared sentences which contained segments and important and contrastive prosodic features which are known to be difficult for L2-learners of Swedish (Bannert, 2004). For this study the focus was on sentences containing minimal pairs contrasted by quantity relations and word stress pattern as well as sentences with compound words or words with a derivational morpheme. In general mispronunciation of segmental phonemes is observed, but only those related to stress and quantity issues were examined in this study.

Participants

Six L1 speakers of Kurdish participated in this study: two male and four female speakers with ages between 36-57 years old when recorded. They had lived in Sweden between 17-25 years, four of them longer than 20 years. Sorani Kurdish is their native language. Four speakers were from Iraq and one is from Iran. All participants worked as interpreters in Sweden and participated in a weekend course about Swedish prosody. They all agreed to participate in a research study. The recordings were done the first day of the course and were entered into a computer using the program Praat (http//www.fon.hum.uva.nl/praat) and headset equipment. Recordings of the same sentences produced by one Swedish male speaker of the standard variety were used for comparison.

Analyses

The stress pattern in the participants' first language, Kurdish, has an impact on learning a second language with another stress pattern. Many of the minimal pairs with a quantitative contrast also seemed to be hard to manage. Four of the speakers pronounce the words with an incorrect vowel quantity in all occurrences. It works in both ways, a long vowel is pronounced as a short vowel in some words and a short vowel is pronounced as a long vowel in other words. The spelling, where double consonants of the same type signal that the preceding vowel is short, does not seem to signal the pronunciation to the speakers. It might be the case that the L2-speakers cannot make use of this hint, which is the major rule in Swedish.

In the read sentences the vowel quantity contrast was present in minimal pairs like *vägen – väggen* [ˈvɛːgən – ˈvɛgːən] (the road – the wall), *granen – grannen* [ˈgrɒːnən – ˈgranːən] (the Christmas tree – the neighbor), *vila – villa* [ˈviːla – ˈvɪlːa] (rest – house), busar – bussar [ˈbʉːsar – ˈbɵsar] (hooligans – buses). The sentences were also constructed with a semantic clue, e.g., *Granen är vackrast med glitter* (the Christmas tree is beautiful with tinsel) and *Man ska vara vän med grannen* (You should become friend with your neighbor). A measurement of the duration of the a-vowel and the following /n/ in the two words shows that there is hardly any difference between the words pronounced by the Kurdish speaker, but for the Swedish speaker there is a difference in both vowel and the following consonant duration in the two words, see Table 1. It is interesting that there is a great difference between the two phonemes in the word *granen* for the Swedish speaker, but almost the same duration in the word *grannen*. However, the quality of the a-vowel is different and probably made more sense for a native listener. This is also shown in the acoustic analysis from Praat, see Figure 1 (Swedish speaker) and 2 (Kurdish L1-speaker). This result conforms the auditory impression when listening to the recordings.

Table 1
Duration in % of the two phonemes /a/ and /n/ for the Swedish and the Kurdish speaker in one of the minimal pairs.

Key words	Phonemes	Swedish speaker	Kurdish speaker
granen	/a/	28%	25%
[ˈgrɒːnən]	/n/	16%	18%
grannen	/a/	21%	24%
[ˈgranːən]			

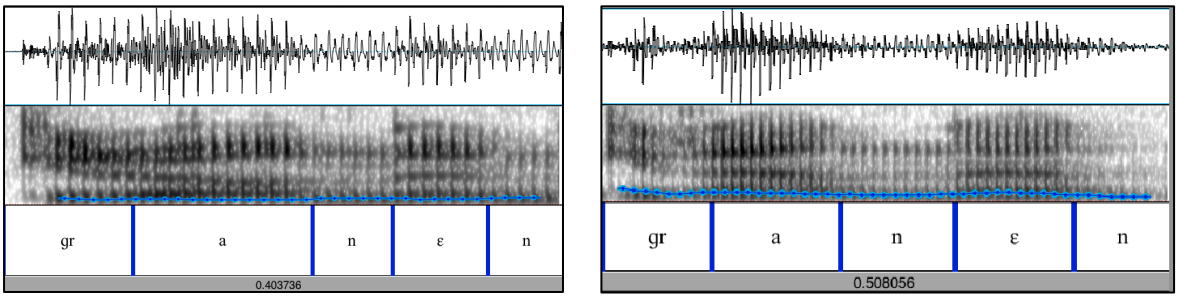


Fig. 1. A Swedish male speaker; *granen* [ˈgrɒːnən] to the left, *grannen* [ˈgranːən] to the right.

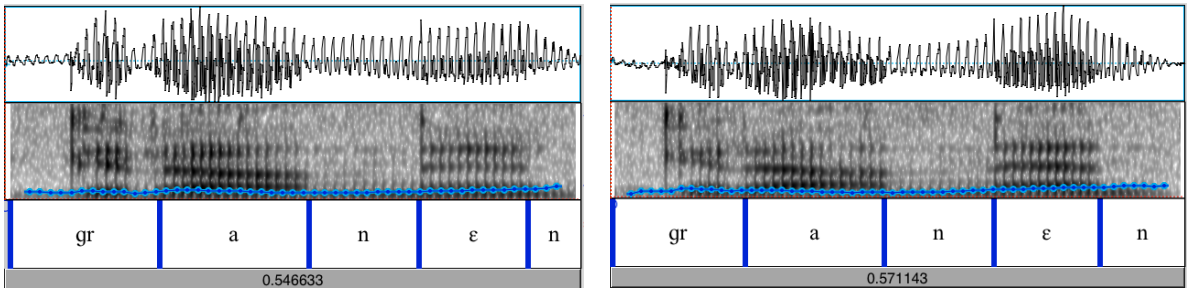


Fig. 2. A Kurdish L1 male speaker; *granen* [ˈgraːnən] to the left, *grannen* [ˈgranːən] to the right.

Three minimal pairs with a contrastive word stress pattern were analyzed in the read sentences. Only one of the minimal pairs was pronounced correctly by all speakers but at least three of the speakers did not change the vowel quality for the two a-vowels in the words spelled *banan* [ˈbɒːnan – baˈnɒːn] (the course – banana) even though they changed the vowel quantity. The same pattern concerning the pronunciation of the a-vowel is shown in the minimal pair *kallas* – *kalas* [ˈkalːas – kaˈlɒːs] (be called – party). It is always the vowel quality for the short a-vowel [a] that is pronounced. One explanation for that is that Kurdish only has this a-vowel quality, but as both a short and a long variety (Rahimpour & Dovaise, 2011). In general, the L2 speakers stressed the last syllable in Swedish, disregarding the variation of stress placement. This is probably a transfer from their first language.

The compound words and words with derivational morphemes seemed to be hard for the speakers, especially for two of those who always place the stress on the last element, like in Kurdish. They had both lived in Sweden for 25 years, but still had problems with the stress pattern for compound words. The other four speakers had a few misplacements each. In Table 2, four examples from the recordings are shown. For these words it was not the syllable duration that seemed difficult, only the stress pattern.

Table 2
Derivation and compound words from the readings

Target words	Translation	Swedish speaker	Kurdish speaker
avlyssna	listen to	[ˈɒːvˌlysna]	[ɒːvˈlysna]
nymålad	recently painted	[ˈnyːˌmoːlad]	[nyːˈmoːlad]
telefonnummer	phone number	[tɛlɛˈfoːˌnɵmər]	[tɛlɛfoːˈnɵmər]
kaffebricka	coffee tray	[ˈkafɛˌbrɪka]	[kafɛˈbrɪka]

One of the female speakers who had lived in Sweden for 17 years was an exception in this analysis. She almost managed to make the prosodic distinction in the minimal pairs, both for vowel quantity and quality, for word stress patterns in isolated words as well as in derived

there was another prosodic feature, namely the Swedish word accent, that seemed to be hard for her. This was not the goal for this investigation but notable because the distinction between the two word accents is often hard for L2 learners of Swedish.

Implications for L2 learners and teachers

Since prosodic features are very important and distinctive in Swedish, it is of importance that second language learners, and teachers have knowledge about them. Speaking with a wrong stress pattern can cause communication errors and problems in a dialogue. The relation between stress and quantity in the syllable structure is of importance, as well as the difference in vowel quality, depending on the quantity. The findings, when measuring the duration in the minimal pair *granen – grannen* (see Table 1 and Figure 1), for the Swedish speaker show the importance of the vowel quality as a clue for a native listener. There are rules for word stress patterns depending on the morphology that should be pointed out. Some derivation morphemes have to be stressed, others do not. Careful listening to Swedish words spoken by native speakers, explanation and guidelines for a correct stress pattern as well as transcription and recordings of the learner's own speech might be a successful pedagogical method to reach awareness of the differences in prosodic minimal pairs (Cauldwell, 2013).

CONCLUSIONS

The prosodic patterns concerning syllable duration and stress placement are different in Swedish and Kurdish. In Swedish, there is a complementary length distinction in stressed syllables and they are produced with a slightly higher intensity and pitch compared to other syllables in the word or the phrase. The stress placement can change depending on the morphology, except for compound words with main stress on the first element and a secondary stress on the last element. In Kurdish on the other hand, stress is not correlated to length in the syllable and stress is usually on the last syllable or last element in a compound. The same stress pattern is found in this study for L2-Swedish. The results indicate that the Swedish stress pattern is confusing for some of the speakers and they place stress at the end of words, as they do in their L1. In some cases the misplacement of stress changes the meaning of the word in question. In those cases where misplacement does not change the meaning, it still can cause communication problems. The theory and earlier research about prosodic transfer between L1 and L2 (Chun, 2002; Mennen, 2006) is thus confirmed. It has to be pointed out that none of the informants in this study live in any of those dialectal areas where stress in compounds can be placed at the last element, e.g. the northern part of Sweden.

Concerning the vowel quantity contrast in stressed syllables of minimal pairs it is important to make a difference both in quantity and quality, especially for the a-vowel, in Swedish. The auditory and acoustic analyses and comparison of the Swedish and the Kurdish speaker presented in Table 1 indicate that both the quantity of the phonemes and the vowel quality is of importance. The Kurdish speaker did not make a clear distinction between the duration of the most important phonemes in the two words and used the same a-vowel throughout, in the minimal pairs in this study.

The speakers in this study learned Swedish after the age of 17, i.e. as adult learners. They are interpreters between Kurdish and Swedish and are highly motivated to acquire Swedish on a high level of proficiency. Their accent is intelligible; however, a better control over the

placement of word stress, would facilitate the comprehension of their Swedish. One of the informants is more or less fluent with an almost correct Swedish prosody, which might be an indication that it is possible to acquire the Swedish prosody for Kurdish L1-speakers regardless of the age of onset and despite the differing word stress patterns between the two languages.

ACKNOWLEDGEMENTS

The research was partly supported by Linnaeus University, Sweden.

ABOUT THE AUTHORS

Elisabeth Zetterholm is a senior lecturer and researcher in Swedish as a second language at Linnaeus University, Sweden. Her current research interests focus on pronunciation in second language acquisition, comprising segmental and prosodic issues. When teaching, most of her students are prospective teachers of Swedish as a second language and therefore she wants to develop the teaching methodology. She received her PhD in phonetics at Lund University, Sweden (2003). Her thesis and postdoctoral research were about voice imitation, with the focus on speaker identification and individual features in voice and speech.

Mechtild Tronnier is a senior lecturer and researcher in phonetics at Lund university, Sweden. Her current research interests focuses foreign accent phenomena and pronunciation issues in second language acquisition, comprising segmental and prosodic issues. She teaches phonetics to speech therapy/pathology students, prospective teachers of Swedish as a second language and students in linguistics on undergraduate, master and graduate level. Previous research includes phonetic questions on prosody, nasality, children's voices and phonetics in second language learning and teaching.

Elisabeth Zetterholm, PhD
Department of Swedish
Linnaeus University
SE-351 95 Växjö
Sweden
Tel: +46 70 5354183
Email: elisabeth.zetterholm@lnu.se

Mechtild Tronnier, PhD
Department of language and literature
Lund University
SE-221 00 Lund
Sweden
Tel: +46 46 222 8440
mechtild.tronnier@ling.lu.se

REFERENCES

Abrahamsson, N. (2009). *Andraspråksinlärning*. Studentlitteratur.

Bannert, R. (2004). *På väg mot svenskt uttal*. Studentlitteratur.

Bruce, G. (2012). *Allmän och svensk prosodi*. Studentlitteratur.

Cauldwell, R. (2013). *Phonology for listening. Teaching the stream of speech*. Speech in Action.

Chun, D.M. (2002). *Discourse Intonation in L2. From Theory and Research to Practice*. John Benjamins Publishing Company.

Hassanpour, A. (2012). The indivisibility of the nation and it linguistic divisions. In: *International Journal of the Sociology of Language* 2012, Issue 217:1-18.

Mennen, I. (2006). Phonological and phonetic influences in non-native intonation. In: Trouvain, J. & Gut, U. (eds.) *Non-Native Prosody. Phonetic Description and Teaching Practice*. Trends in Linguistics. Studies and Monographs 186. Mouton de Gruyter.

Munro, M. (2008). Foreign accent and speech intelligibility. In: J.G. Hansen Edwards & M.L. Zampini (eds.) *Phonology and Second Language Acquisition*. Amsterdam: John Benjamins Publishing Company.

Praat: www.fon.hum.uva.nl/praat

Rahimpour, M. & Dovaise M.S. (2011). A Phonological contrastive Analysis of Kurdish and English. In: *International Journal of English Linguistics*. Vol. 1, No 2; Sept 2011.

Riad, T. (2013). *The phonology of Swedish*. Oxford University Press.

Tronnier, M. & Zetterholm, E. (2011). New Foreign Accents in Swedish. In: *Proceedings ICPhS 2011*, HongKong 2011:2018-2021.

www.scb.se/statistik/_publikationer/UF0524_2012A01_BR_05_UF01BR1201.pdf

Lan, Y. (2014). Detecting L2 speech deviations by a communicative experiment procedure: Cantonese speakers' realizations of English /r/. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 75-86). Ames, IA: Iowa State University.

DETECTING L2 SPEECH DEVIATIONS BY A COMMUNICATIVE EXPERIMENT PROCEDURE: CANTONESE SPEAKERS' REALIZATIONS OF ENGLISH /r/

Yizhou Lan, City University of Hong Kong

One purpose of connecting L2 speech research and L2 teaching and learning is to help students learn to communicate more effectively in spontaneous speech. However, experiments with L1 and L2 speech production have long used carefully controlled reading procedures. Such procedures may not predict performance in real-time communication. In order to compare reading and self-generated speech, two experiments examined how native English speakers' ratings of native-likeness for Hong Kong English speakers were affected by experimental procedure. Participants were 8 advanced Cantonese speakers of English pronouncing real *r- or -r-* words. In experiment 1, participants read the stimuli carefully in a carrier sentence; in experiment 2, participants were told to make up a story out of the same stimuli used in experiment 1. Results showed that in experiment 2, more errors and types of errors were noticed by native English speakers. Results imply that gearing the procedure to a more functionally-loaded one will more fairly evaluate actual speech performance.

INTRODUCTION

L2 speech acquisition is partly determined by linguistic experience (Flege, Schirru, & MacKay, 2003), and other distributive learning models also support this idea from both L1 (Pierrehumbert, 2003; Saffran, Aslin, & Newport, 1996) and L2 (Best, 1995; Kuhl, 2000) data in speech perception. The types of errors in L2 production can also be attributed to the speakers' language experience. However, other studies have proposed that development errors also exist (Ellis, 1994). They postulate that the sequence in which L2 errors appear echo that of native speakers' speech errors along similar developmental stages. However, most experiments looking at L2 production patterns have been done in strictly controlled settings with participants carefully reading words, sentences or paragraphs. The present study looks at a different experimental paradigm of L2 pronunciation learning, with eight Cantonese speakers (NC) pronouncing the phoneme /r/ in L2 English.

The phonemes which Native English (NE) speakers may find difficult or late in development (in this case, /r/) may also be difficult for NC speakers because more competing cues are present and which cue attracts L2 learners' attention can be highly variable (Davidson, 2006). One representative is the sound /r/ in production because it has both gestures on the tongue tip and tongue body (Browman & Goldstein, 1992). However, according to Speech Learning Model (SLM, see Flege 1987), speech sounds that are very different should be learned more easily by L2 learners whose native language does not include a similar phonemic or allophonic category.

Therefore, testing this sound may help see if the production of /r/ is consistent with the SLM or with a developmental difficulty model.

Towards an alternative method

One purpose of connecting L2 speech research and L2 teaching and learning may be to help improving L2 learners and listeners' intelligibility and facilitate smoother communication (Derwing & Munro, 2003). However, the reliability and validity of experiments of L1 and L2 speech production have long been accredited to the convenience of controlled careful reading procedure of reading single words (Gonzalez-Bueno & Quintana-Lara, 2012), single sentences with the carrier word inside (e.g., Aoyama, Flege, Guion, Akhane-Yamada & Yamada, 2004; Best & Taylor, 2007; Chan, 2006; Flege, 1987, 1995; Hung, 2002, to name just a few), or reading a passage (Lan & Oh, 2012). These studies often discuss the communicative competence of the learners by the results of careful reading. Nevertheless, in a communicative language teaching context, the careful-reading task faces two challenges. First, does careful reading represent the performance in real-time communication? Second, in controlling other elements of linguistic processing (as in reading, not self-generation of language), are we isolating pronunciation as an independent modular language process? Although spontaneous reading has been advocated as a means of pedagogical application, few studies on speech learning have used this approach (Chan, 2010; Sachet, 2013). In Chan's study, the spontaneous speech method was one of the three procedures (sentence reading, paragraph reading, and spontaneous speech) and the results were not analyzed by acoustic measures.

Cantonese L2 perception and production of English /r/

According to Matthews and Yip (1994), /r/ is not in present in the Cantonese consonant inventory. According to Flege (1995), the absence of the /r/ in Cantonese may lead to easier acquisition and thus fewer errors, because a new sound is easier to acquire than the similar sound. L2 sounds similar to L1 ones are more easily assimilated to L1, and thus harder to acquire. Chan (2006), in support of Flege's assertion, pointed out that over 80% of the /r/-initial tokens were pronounced correctly in production tasks. But is this sound really so easy for NC speakers? The complexity of this sound (Chan, 2006) indicate that acquisition of /r/ is not straightforward. Previous studies have shown that /r/ was often pronounced as [w], as depicted by high F3 values, by NC speakers (Hung, 2002). Moreover, in clusters, /r/ tends to be phonologically deleted (Chan, 2006).

Specifically, in previous studies, we found that /r/ is phonetically reduced (not totally substituted to [w]) in reading tasks. According to Lan and Oh (2012) as well as the interview with students prior to the experiment, the error types were not limited to those derived of previous studies. In addition, Cantonese learners in the current study would realize /r/ as [l] or [t]. Since /r/ is perceptually most similar to /w/ (Chan, 2006; Hung, 2002; Lan & Oh 2012), the production of [l] and [t] are phonologically surprising. This is not reported by previous studies using reading procedures. This suggests the necessity of a new way to mine out more error types and approximate the real situation of pronunciation learning problems.

METHOD

The study examines acoustic properties and error types of production of /r/ by experienced Cantonese learners of English in two different procedures. The experiment uses a comparative analysis between careful reading and spontaneous speech.

Participants

-
- NC Participants were eight adults working as administrative staff at City University of Hong Kong (4 females, 4 males, mean age=27.5). They all used English as their working language with at least 18 years of English learning experience. None of them had exposure to other foreign languages except English. All participants were right-handed with no reported hearing or motor-control defects. They did not have prior exposure to musical training. Control speakers were two NE speakers (1 female and 1 male, mean age=26.5) from California, U.S.

Materials

Two experiments used the same stimuli set as materials but with different collection procedures. To let their speech contain as many [r] tokens as possible, 18 words in CVC structure were designed to be tested in both experiments. Five words containing /r/-initials varying in five vowel contexts of /i, æ, u, ʌ, and ɔ/, and three words containing [r] clusters were used as /r/-targets. Five /w/-initial words were added as controls because of previously reported /r/-/w/ confusion by NC speakers. Moreover, five other CVC words were inserted to the list as distractors. The wordlist was *root, rob, read, rat, rub, print, train, cream, wok, wear, weak, wide, wake*, *cheap, dark, goat, cop and think*, which are all meaningful and frequent words and were known by the subjects.

Procedure

In experiment 1, we tested the /r/ productions in sentence reading. Carrier sentences of "Now I say ____." were instructed to be read by participants. Randomized stimuli words (18 words excluding fillers) were inserted to the blank and presented to participants. The number of tokens was 13 words × 10 participants (including NC and NE) = 130.

The procedure in experiment 2 was not strictly controlled and took the form of free speech in a laboratory setting. In a pilot study, we let participants make up two stories after five minutes of preparation, with both stories including all the words printed on a word-list containing the randomized stimuli words as in experiment 1. However, two pilot subjects thought the task was too difficult to complete. So we broke the randomized word-list down into four smaller word sets. At the start of the second experiment, participants were told we were testing fluency in spoken English.

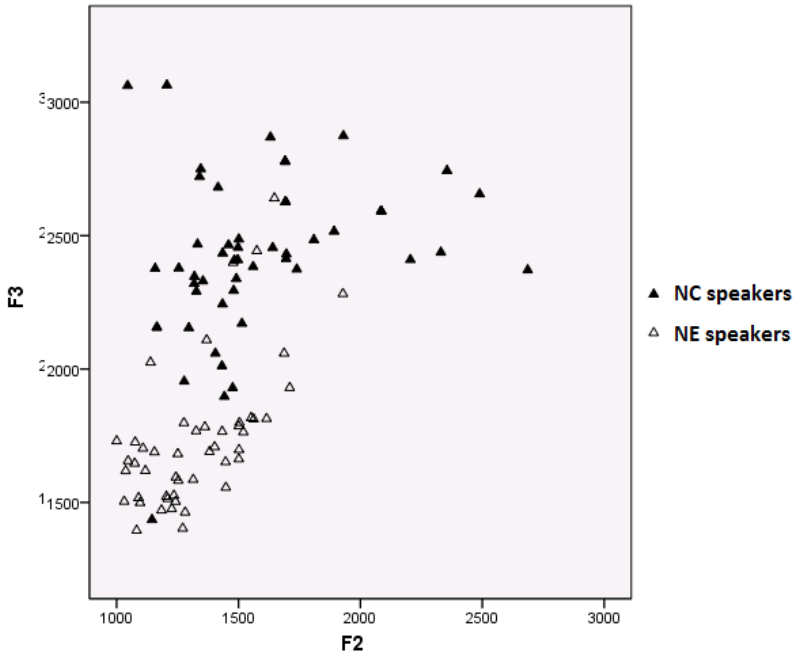
In experiment 2, altogether 101 usable tokens (including /r/-initials, /w/-initials and /r/-cluster initials) were collected from five Cantonese participants' productions and 48 tokens from the native English participant's productions (101+48=149 tokens). Stimuli tokens for each

participant differed with some overlap. Only five of eight NC participants successfully recorded in Experiment 2 in the sound booth because other three promised to make a spontaneous speech prior to the experiment but refused to do so on spot.

The productions of target words (including /r/-initials, /r/-clusters and /w/-initials as control) in both experiment 1 and 2 were extracted from the sentences and segmented as phonemes within those words. The /r/ parts of the productions, defined as the section from the beginning of voicing to the steady state of vowel, were examined for its second formant (F2), third formant (F3) and general audible perception judgment by NE speakers. Another two Native English listeners with standard American English accent who were not involved in the production experiment and a phonetically trained Chinese speaker worked together to rate the productions in its nativeness and denote the types of mis-pronunciations by Cantonese participants' English productions in both experiments.

RESULTS

Figure 1 exhibits two scatter plots with the upper one displaying the F2 and F3 formant space in /r/ produced by NE and NC speakers in the reading task and the lower one displaying the same space of the spontaneous task.. From both figures we can see that as expected, NE speakers show smaller F2 and F3 values in distribution, which is an acoustic property of /r/. Regardless of task, we see a general tendency of Cantonese speakers to mispronounce /r/ by increasing the F3.



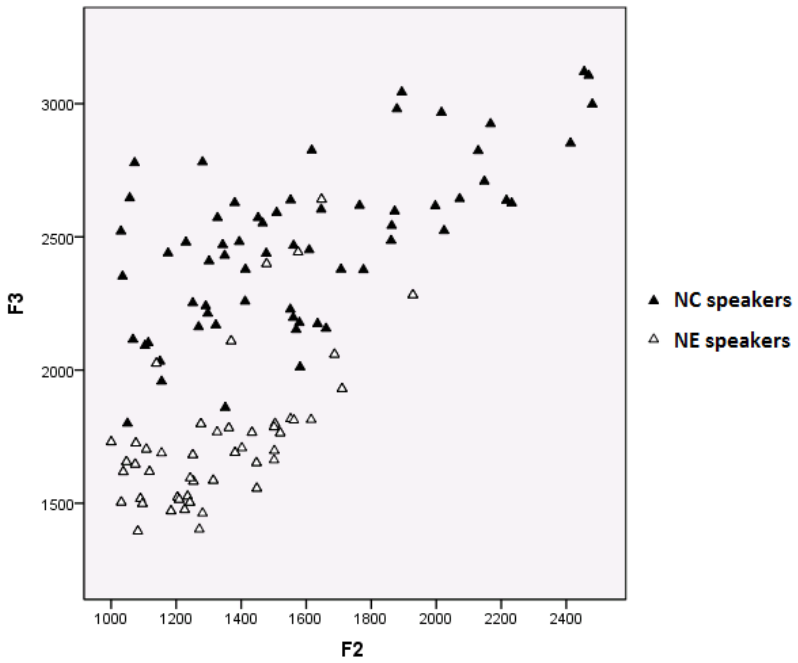


Figure 1. Above: Scatter plot of NC and NE speakers' F2 (horizontal axis) and F3 values (vertical axis) in the reading task. Below: Same F2 and F3 in the spontaneous speech task. NC speakers showed higher F3 in both tasks.

Formant values for words with /r/ in two experiments

The data for NE speakers' spontaneous speech were used in both experiments because there was not much difference for the NE speaker's productions in different settings. For NE speakers, the mean formant values were F2=1314 and F3=1702 (N=46, analyzable tokens). A t-test showed that the difference between NC and NE participants in experiment 1 was significant (F2: t=4.204, df=92, p<.0001; F3: t=13.256, df=92, p<.00001).

On average, the formant values in the careful reading task (experiment 1) for /r/ initial and /r/ cluster words by NC speakers were F2=1573 and F3=2410 (N=48, analyzable tokens). Inter-subject difference was not significant (F2: F(2, 45)=1.300, p=.283; F3: F(2, 45)=.293; p=.747). The difference of /r/ in word-initial or in cluster was not significant as well (F2: t=-.586, df=46, p=.560; F3: t=-.600, df=46, p=.551).

For the spontaneous speech task in experiment 2, the mean formant values for NC speakers were F2=1492 and F3=2232 (N=84). NC participants did not show a significant inter-subject difference for F2 (F(2, 81)=2.528, p=.087). However, the inter-subject difference for F3 was significant (F(2, 81)=8.671, p<.001). The difference of /r/ in two word types was not significant in F2 (t=-2.243, df=82, p=.28). However, the difference of /r/ in initial or cluster positions was significant for F3 (t=-5.589, df=82, p<.001). Again, a t-test showed that the difference between NC and NE was significant (F2: t=2.884, df=126, p<.05; F3: t=7.530, df=126, p<.0001).

The comparison between two experiments was insignificant for F2 (t=1.232, df=130, p=.220), but significant for F3 with the production in experiment 1 having a higher F3 (t=2.637, df=130,

p<.01). A pictorial representation of the results was depicted in Figure 2 below (error bars at 95% Cl).

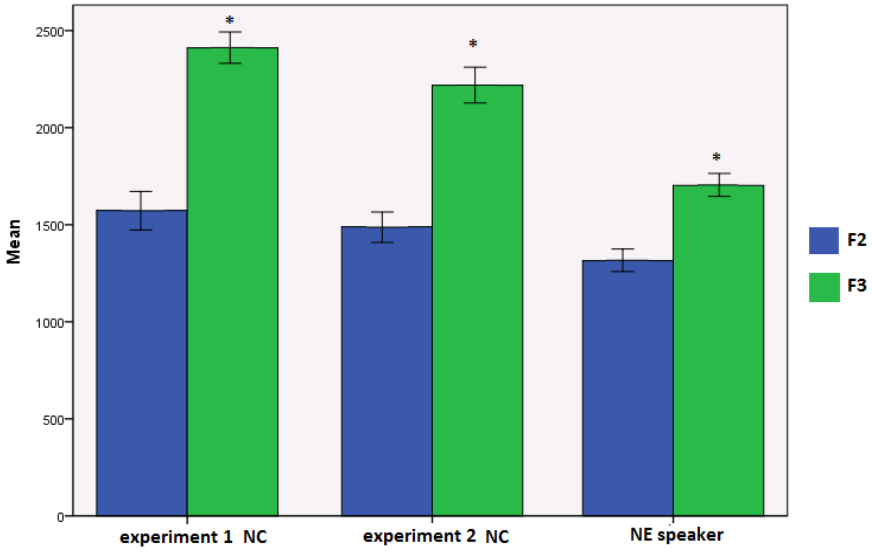


Figure 2. Comparison of F2 and F3 values of NC speakers' (separately in two experiments) and NE speaker's production (in experiment 2).

Comparison between /r/ and /w/ in two experiments

A comparison of initial /r/, cluster /r/ and /w/ production was done for NC speakers in both experiments 1 and 2 to see how the difference between /r/ and /w/ was realized in production. Surprisingly, in experiment 1, the difference between the three entries was not significant for both F2 and F3 [F(2, 51)=.886, p=.427; F3: F(2, 51)=2.103, p=.133]. Speakers did not distinguish between /r/ and /w/ in production. However, in experiment 2, the difference was not significant for F2 [F(2, 91)=2.601, p=.08] but was for F3 [F(2, 91)=14.537, p<.001]. Tukey's post-hoc tests revealed that the significance lay in the difference between cluster /r/, which has an even higher F3 than the average of /w/ [md=-490.481, Std.E=91.053, p<.001]. A pictorial representation of the results was depicted in the figure below (error bars at 95% Cl).

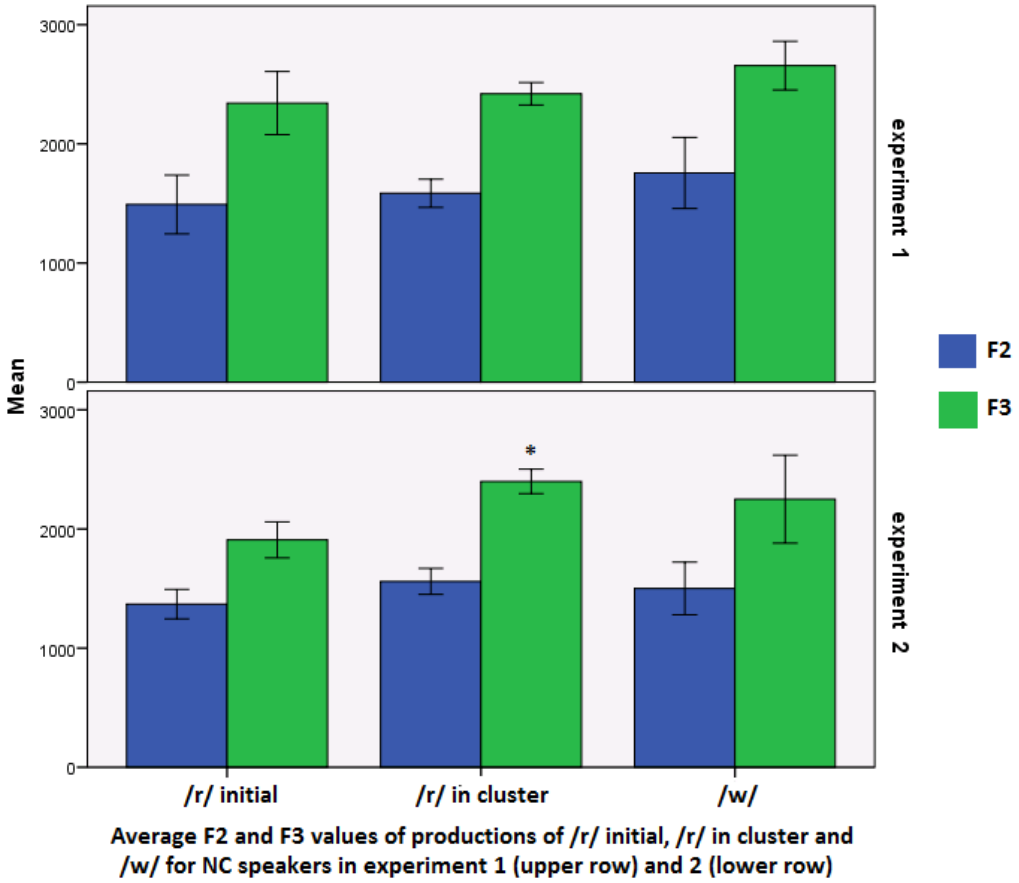


Figure 3. Comparison of F2 and F3 word type for NC speakers' production in two experiments.

Native speaker perception

Three listeners, including two NE speakers and one Chinese phonetician, were asked to pick which productions were more accented, similar to the way Flege et al. (2003) assessed accentedness. Words pronounced as strongly accented were labeled "1" and mildly or unaccented sounds were labeled "0". Due to technical problems with the computer when the listeners were doing the perception test, only the data for experiment 2 were collected. The inter-rater difference is not significant. A correlation between formant values and native speaker's perception was done. Spearman's correlation test showed that F2 was not significantly correlated with NE perception (r=.03, p=.324). However, F3 was correlated with NE perception in near-significance (r=-.15, p=.078). Such a correlation was negative, indicating that the lower the F3 was, the better chance it could be perceived as a good token for /r/. This suggests that using F3 as an indicator for native-like English /r/ production has some perceptual basis.

Error patterns

Only r-deletion and [w]-substitution were found in the first experiment. However, three more error patterns were found in experiment 2. Firstly, bidirectional confusion of /r/ and /l/ as well as /r/ and /w/ was found. Secondly, hypercorrection of inserting /r/ and /l/ in non-cluster words,

such as *pay* as *play*, or *big* as *brig*, was found. Finally, affrication of /r/ in the /r/-initial was present as well.

The most common mistake was the complete omission of [r] in the production of *tree*. The second type of error is the substitution of [w] in *Troy*. Spectrograms of these two words are shown below in Figure 4. In *tree* on the left panel, the vowel part is followed right after the noise part with no visible F3 variation (Please refer to the arrow in on the left panel). In *troy* on the right panel, a signature F1-F2 nearing which indicates [w], instead of F2-F3 nearing indicating [r] is represented (shown by the arrow in on the right panel).

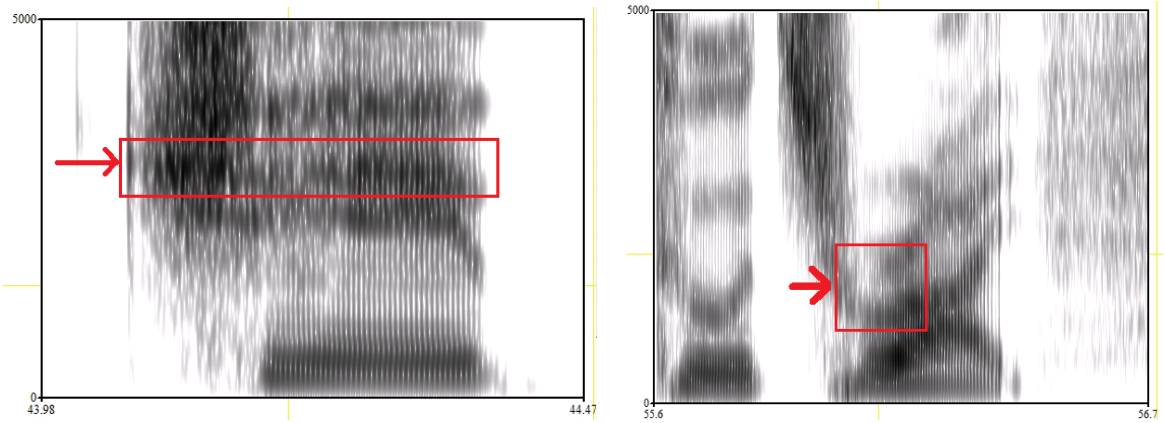


Figure 4. Spectrogram examples of error types 1 and 2: deletion (left) and substitution (right).

Error types 3 and 4 were not found previously in studies of careful reading. The third type is the substitution of /l/ in this type of error: *pray* is pronounced as *play* twice in the third speaker's pronunciation. Similarly, *clean* was pronounced as /krin/. The spectrogram in Figure 5 on the left is the production of *play* . The visible rising F3 indicates the presence of /r/ (shown by the arrow in on the left panel). Moreover, /r/ and /w/ were also found to be acoustically similar . This resulted, in some cases, that the F3 value in /w/ tokens were lower than that of /r/ ones because some productions with low F3 value in /r/ were realized with higher F3 value as in /w/, and vice versa. This might be related to the high variation of F3 in the second experiment, though it may not be a direct cause of it.

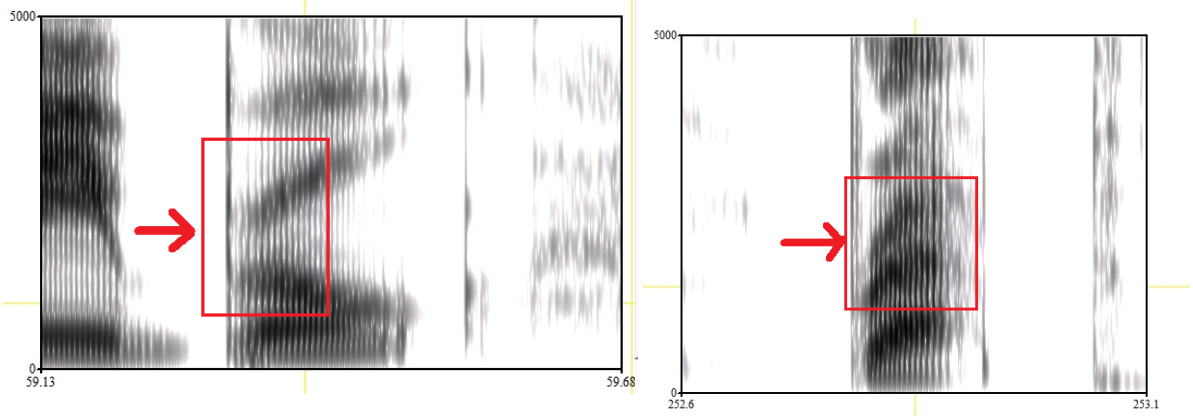


Figure 5. Spectrogram examples of Error types 3 and 4: /l/-substitution (left) and hypercorrection (right)

The fourth type of error was a hypercorrection from non-/r/ CVC sounds to clustered CrVC sounds. The above spectrogram on the right shows *big* with /r/ inserted between the consonant and vowel because of the apparent F3 rise. This was not a slip of tongue because /r/ insertion occurred seven times in the five speakers' productions, while native speakers of English did not depict any of such insertion in non-/r/ words.

DISCUSSION

Differences between two procedures and its implications

NE speaker perception results, with ratings of nativeness and the F3 value being negatively correlated, has allowed us to use F3 as the perceptual factor that could be used to examine nativeness of production of /r/. In both experiments, although both /r/-deletion and [w]-substitution were found in the production data , not all tokens had had their /r/ phonologically deleted or substituted. Phonological substitution of /r/ by [w] should entail the acoustic realization of /r/ to have a similar distinctive feature of [w]. However, the actual production results did not show a universal high F3 as in /w/ produced by the same group of speakers. Therefore, it is proposed that the category of /r/ is not phonologically mapped by L1 categories, but a new category established on the basis of both L1 and L2.

In the comparison of formant values between F2 and F3, the results showed, quite unexpectedly, that in the performance of /r/ production in terms of F3 was even more non-native-like in the reading task than in spontaneous speech, although the discrepancy was not significant. This indicates that the *degree* of pronunciation deviations, demonstrated by the F3 values in production and proven with its connection by NE perception, did not vary much even when NC participants were faced with a cognitive load to produce sounds or with reading.

However, the individual differences were significant in the second experiment, leaving the impression that in spontaneous speech, the quality of the sound is more unstable within the /r/ sounds. Compared with /r/ and /w/, the cluster /r/ was utilized as a sound with different acoustic distinction to differentiate from /w/, although we were not sure if the difference was perceptual as well. Since the gestures for C-/r/ clusters involved both tongue tip, tongue dorsum and lip, and arranged in short time with considerable overlapping, the C-/r/ clusters can be seen as complex in gesture (Browman & Goldstein, 1992; Lan & Oh, 2012). When processing these clusters, if the learners were in cognitively more loaded situations such as spontaneous speech when they do not have enough cognitive resources to focus on pronunciation, the variation of F3 may be higher in such a condition.

Implications for L2 teaching and research

Morley (1994) pointed out that heavy cognitive loads in independent speech tasks will lead to more pronunciation errors and hence it is necessary to incorporate guided practice and spontaneous speech together in an integrated curriculum. In her curriculum guidelines for instructional planning of pronunciation courses, a practice mode that moves from dependent practice through guided practice to independent practice was introduced. The last one is

represented by extemporaneous speech. The study takes one step further to urge for an update of methods in pedagogy-oriented research as well.
Larsen-Freeman (1997) referred to chaos theory to explain SLA in general, that one simple rule of difference may generate various unpredictable patterns of production. In the aspect of speech production, we could also see that the very limited feature components in one phoneme can result in various perception and production errors either due to lack of attention of one or a combination of more of the features, by interventions of the L1 category directly, or by other non-phonological habit-formation such as sociolinguistic hypercorrection (Chan, 2006), insertion due to lexical influence (Setter, 2008), and lack of motor control (Browman and Goldstein 1992; Davidson, 2006).

Therefore, in looking at L2 speech production, both linguistic factors such as the influence of L1 phonology on L2, and learners' affective factors should be both considered. The latter component is discussed in the following section. Since the factors that were neglected in traditional studies regard the L2 system as merely a mixture of L1 factors and L2 factors as well as phonological transfer rules, we need to introduce procedures with less control but more inclusion of the learner variables, even at the expense of controlling variables exhaustively.

Learner variables

One difficulty to carry out the study was the affective factors of participants as learners, especially the attitudes of learners towards their accentedness of pronunciation. Just as Derwing (2003) reveals, L2 learners are aware of their general pronunciation deficiency and appear shy and unwilling to have their voice recorded, especially in spontaneous tasks, where only five out of eight participants successfully finished the task. Even in one participant's recording (he went through the tasks successfully, of course), whisperings, in Cantonese, of "this is too difficult to me, tell me how to do it" were found. However, this attitude in turn led to more pronunciation problems because genuine production tasks (facing a recorder or communicating with a native speaker) may witness more pronunciation errors because of lack of self-monitoring and stutter/hyper-caution/avoidance. One piece of evidence in the production is the hypercorrection errors, such as pronouncing *pay* as *play*, or *big* as *brig*. The fear to mis-pronunce "Cr-" clusters has led to the adding of "-r" color productions even in cases "r" is not presented. This is similar to other parts of grammar in Cantonese English interlanguage: in Cantonese English, the plural marker (-s) and past tense marker (-ed) are often misused in English sentences to avoid mistakes (Chan, 2006).

As for motivation, these experienced learners in Hong Kong were very keen on acculturating to the English language community from the perspective of pronunciation. In an interview after the experiments, they confessed that they never resisted changing their pronunciation and never wanted to keep the Cantonese identity in spoken English. However, this view of acculturation has a disadvantage – they fear speak and being discovered to have imperfect pronunciation.

The study examined the effect of experimental procedure on production test results and proposed a novel way of examining L2 speech production. Results partly agreed to the prediction that multiple error types would occur in experiment 2, but the quality of sounds in experiment 2 were

more native-like in terms of formant values, though they were more variable than that of experiment 1.

ACKNOWLEDGEMENTS

The author thanks Dr. Li Bin for her suggestions to an earlier draft of this paper. The author also thanks Karen Chow for helping recruiting participants.

ABOUT THE AUTHOR

Yizhou Lan is a Ph.D. Candidate and Teaching Assistant in linguistics at City University of Hong Kong. His research interests include second language acquisition, speech perception and production, motor theory, speech learning models, and pronunciation teaching and learning. His current project is on Cantonese and Mandarin advanced learners' production and perception of English sounds, and intends to add more cognitive features to current speech learning models by experimentation. His publications include *Proceedings of 4th International Symposium on Tonal Aspects of Languages* (published by ISCA), *Proceedings of PACLIC 27*; *Proceedings of IEEE ICSPCC 2013*, and *Journal of the Acoustical Society of America*, He expects to graduate in Summer 2015.

REFERENCES

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese/r/and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233-250.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 171-204.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. In O.-S. Bohn & M. J. Munro (eds.). *Language experience in second language speech learning: In honor of James Emil Flege*, 13-34.

Browman, C.P. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica, 49,* 155-180.

Chan, Y. (2006). Strategies used by Cantonese speakers in pronouncing English initial consonant clusters: Insights into the interlanguage phonology of Cantonese ESL learners in Hong Kong, *IRAL proceedings, 44,* 331-355.

Chan, A. Y. (2010). An investigation into Cantonese ESL learners' acquisition of English initial consonant clusters. *Linguistics, 48*(1), 99-141.

Davidson, L. (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences, *Journal of Phonetics*, *34,* 104-137.

Derwing, T.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.

Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press, USA.

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

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. In W. Strange (ed.). *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.

Flege, J. E., Schirru, C. & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems, *Speech Communication, 40*, 467-491

Gattegno, Caleb (1963). *Teaching foreign languages in schools: The Silent Way* (1st ed.). Reading, UK: Educational Explorers. Retrieved October 10, 2011.

Gonzalez-Bueno, M. & Quintana-Lara, M. (2012). The teaching of second language pronunciation through processing instruction. *Applied Language Learning, 21* (1&2), 57-81.

Hung, T.N. (2002). Language in contact: Hong Kong English phonology and the influence of Cantonese. In Kirkpatrick, A. (ed.). *Englishes in Asia: Communication, Identity, Power and Education* (191-200). Melbourne: Language Australia.

Kuhl, P. K. (2000). A new view of language acquisition. *Colloquium of PNAS*, pp. 11850–11857.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, *18*(*2*), 141-165.

Lan, Y. & Oh, S. (2012). Comparative L2 English [r]Cluster Production by Cantonese and Mandarin Speakers, paper presented at *Conference on Bilingualism and Comparative Linguistics*, Chinese University of Hong Kong, Hong Kong.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(*1*), 1-86.

Morley, J. (1994). *Pronunciation Pedagogy and Theory: New Views, New Directions*. 1600 Cameron Street, Suite 300, Alexandria, VA 22314.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and speech*, *46*(*2-3*), 115-154.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274* (*5294*), 1926-28 .

Sachet, L. M. (2013), *Phonomotor training for accent modification*, Unpublished Master's Thesis. University of Washington.

Setter, J., (2008). Consonant clusters in Hong Kong English. *World Englishes*, *27* (3), 502–515.

Shoji, S. (2014). Japanese epenthetic vowels: How Japanese speakers pronounce English words. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 87-103). Ames, IA: Iowa State University.

JAPANESE EPENTHETIC VOWELS: HOW JAPANESE SPEAKERS PRONOUNCE ENGLISH WORDS

Shinichi Shoji, University of South Carolina

The primary objective of this study was to examine which epenthetic vowels Japanese speakers utilize in pronouncing English words. An analysis of Japanese loanwords from English indicates that Japanese speakers utilize [ɯ] as the context-free default epenthetic vowel, and [o] and [i] as the context-dependent epenthetic vowels. In the framework of the Optimality Theory (OT), I propose that these epentheses are explained with the constraints in ranking of, *COMPLEX, CODACOND >> MAX-IO >> PALATAL-FRONT, SYLLABLE INVENTORY STRUCTURE (SIS), IDENT-IO >> *ROUND, *LOW >> *FRONT >> HIGH >> BACK >> DEP-I. This study claims that native Japanese speakers pronounce English words based on this constraint-ranking.

INTRODUCTION

Foreign words include some sounds or phonological patterns that do not fit the phonetic or phonological constraints in another language. Therefore, the pronunciation of foreign words is "repaired" to "do as little violence as possible" to these constraints (Sapir, 1921, p. 210). This paper specifically discusses epenthetic vowels in Japanese speakers' pronunciation of English words in the framework of the Optimality Theory (OT). This analysis relies on the pronunciations of Japanese loanwords from English because loanwords exhibit how the speakers of the recipient languages modify foreign phonemes and phonotactics. While most of the early studies on Japanese loanwords are limited to explaining partial phenomena, this paper aims to include as many epenthetic patterns as possible in a unified explanation with a single ranking of constraints. This study suggests that this single ranking of constraints explains how native Japanese speakers pronounce English words.

Vowel Epenthesis

The fundamental motivation for epenthetic vowels in loanword adaptation is that recipient languages do not allow codas and consonant clusters. This draws the markedness constraints of NOCODA / CODACOND and *COMPLEX. In order to repair codas and consonant clusters, recipient languages apply deletion or epenthesis based on the faithfulness constraints, DEP-IO or MAX-IO, respectively. According to the preservation principle (Paradis & Lacharité, 1997), epenthesis should be preferred over deletion in order to preserve the input features. An empirical observation shows that a majority of the languages in the world prefer epenthesis rather than deletion (Kang, 2011).

Roughly speaking, epenthetic vowels include context-dependent vowels and context-free default vowels. The context-free default vowel should be as faithful as possible to the empty spot that bears no feature or duration of time. Thus, the epenthesized segment should be the one which is

the least intrusive, the most unmarked, and is perceptually the closest to zero (or silence) in the recipient languages (Hirayama, 2003; Steriade, 2009). In other words, this type of epenthetic vowel should have minimal salience, which would result in the smallest perceptual change between the input and the output.

According to Lehiste (1970), Carr (1999) and Blevins (1995), high vowels are less sonorous and shorter in duration than low vowels. In addition, according to Lombardi (2002), front vowels are more marked than back vowels. Kager (1999) also maintains that [-low, + back, -round] vowels are the most unmarked values for epenthetic vowels. Based on these studies, I propose the constraints of HIGH, *LOW, BACK, *FRONT and *ROUND.

Table 1
Markedness Constraints for Context-free Epenthetic Vowels

Constraints	Functions
HIGH:	Vowels should be high.
*LOW:	Vowels should not be low.
BACK:	Vowels should be back.
*FRONT:	Vowels should not be front.
*ROUND:	Vowels should not be rounded.

The constraints should be ranked in the order that would most accurately generate the epenthetic vowels that actually occur.

A five-vowel system with [a], [e], [i], [o] and [u] is the most common vowel system in world languages. Fijian is one of the languages that have this five-vowel system. Kenstowicz (2003) observes that [i] is the default epenthetic vowel of Fijian, and it is perceptually the closest to zero among the five vowels. However, this is a front vowel. Thus, the constraints about the back feature such as *FRONT and BACK should be lower ranked than the other constraints in order for the constraints to generate [i] as the epenthetic vowel. The optimality of [i] in Fijian is drawn with the ranking below.

Table 2.
*OT Tableau 1. Generation of [i]: HIGH, *LOW, *ROUND >> BACK, *FRONT*

CVC	HIGH	*LOW	*ROUND	BACK	*FRONT
CVCa	*!	*(!)		*	
☞ CVCi				*	*
CVCu			*!		
CVCe	*!			*	*
CVCo	*!		*(!)		

Meanwhile, Kager's (1999) observation of many languages in the world indicates that [ɨ], [ə] and [i] are the most common epenthetic vowels. If we add [ə] to the tableau above, it cannot be optimal because [ə] violates HIGH. Here, we need to place the constraint, HIGH, lower. Also, *FRONT has to be higher than HIGH and BACK in order for [ə] to be optimal.
Table 3.

*OT Tableau 2. Generation of [ə]: *ROUND, *LOW >> *FRONT >> HIGH, BACK*

CVC	*ROUND	*LOW	*FRONT	HIGH	BACK
CVCa		*!		*	*
CVCi			*!		*
CVCu	*!				
CVCe			*!	*	*
CVCo	*!			*	
☞ CVCə				*	*

If we check again the Fijian epenthetic vowel [i] with this constraint-ranking, we can confirm that [i] is successfully generated.

Table 4.
*OT Tableau 3. Generation of [i]: *ROUND, *LOW >> *FRONT >> HIGH, BACK*

CVC	*ROUND	*LOW	*FRONT	HIGH	BACK
CVCa		*!		*	*
☞ CVCi			*		*
CVCu	*!				
CVCe			*	*!	*
CVCo	*!			*	

If all the most common epenthetic vowels in the world, [ɨ], [ə] and [i], are available in a language's vowel inventory, the constraint-ranking above would generate [ɨ] as the optimal epenthetic vowel as shown below.

Table 5.
*OT Tableau 4. Generation of [ɨ]: *ROUND, *LOW >> *FRONT >> HIGH, BACK*

CVC	*ROUND	*LOW	*FRONT	HIGH	BACK
CVCa		*!		*	*
CVCi			*!		*
CVCu	*!				
CVCe			*!	*	*
CVCo	*!			*	
CVCə				*!	*
☞ CVCɨ					*

According to this analysis, among the three most common epenthetic vowels, [ɨ], [ə] and [i], [ɨ] is considered to be the best qualified epenthetic vowel, [ə] the second best, and [i] the third best, although all three are equally common epenthetic vowels. According to Harrison and Kaun (2000), a language has to utilize vowels that exist in its vowel inventories. They express this constraint as IS (INVENTORY STRUCTURE), that is, using a segment that exists in the language's inventory. Thus, it is considered that, in world languages, [ɨ] is not as common as [ə] or [i], and [ə] is not as common as [i]. If a language does not have [ɨ], it utilizes [ə]; if a language does not have [ɨ] or [ə], it utilizes [i]. IS would also explain why other high back unrounded vowels such

as [ɯ] are not as common epenthetic vowels as [ɨ], [ə] and [i]: not many languages have [ɯ] while many of them commonly have [ɨ], [ə] and [i].

Japanese Vowel Epenthesis

Context-free Default Epenthetic Vowel [ɯ]

Japanese does not allow consonant clusters or codas (except [n]), and its loanwords prefer epenthesis to deletion. This draws the following ranking: *COMPLEX, CODACOND >> MAX-IO >> DEP-IO. Japanese has a five-vowel system of [a], [e], [i], [o], and [ɯ] as shown in Figure 1 below. Among those vowels, [ɯ] does not accompany lip protrusion, so it is unrounded. Also, [ɯ] is quite centralized in the Japanese inventory. Japanese [e] and [o] are slightly more centralized than equivalent cardinal vowels.

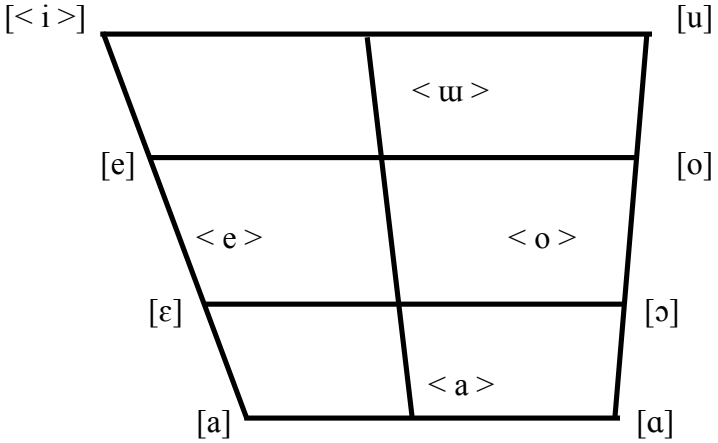


Figure 1. Japanese vowels in < > and cardinal vowels in [ ] (Vance, 2008, p. 54)

In most Japanese loanwords, the epenthetic vowel is [ɯ], which works as the context-free default epenthetic vowel. It is the most unmarked and perceptually least salient among Japanese vowels, as supported by facts that [ɯ] is "the most readily subject to devoicing" and "weakening and elimination" (Lovins, 1975, p. 106; Mori, 1929, p. 58), and that it rarely attracts an accent (Yoshida, 2006). Also, Han's (1962) examination of the duration of Japanese vowels shows that [ɯ] is the shortest in length. Therefore, if we organize the Japanese vowels in order from shortest to longest, we would get the following series: [ɯ], [i], [o], [e], and [a]. The constraint-ranking from the previous section, which generates the default epenthetic vowels, successfully selects [ɯ] as the most optimal candidate as shown below.

Table 6.

*OT Tableau 5. Generation of [ɯ] in Japanese: *ROUND, *LOW >> *FRONT >> HIGH, BACK*

CVC	*ROUND	*LOW	*FRONT	HIGH	BACK
CVCa		*!		*	*
CVCi			*!		*
☞ CVCɯ					*
CVCe			*!	*	*
CVCo	*!			*	

Context-dependent Epenthetic Vowel [i]

Some Japanese loanwords also exhibit the context-dependent epenthetic vowels, [i] and [o]. As for [i], it appears after the palato-alveolar affricates, [*tʃ*] and [*dʒ*], in source words. Examples are shown below.

Table 7.
Loanwords with Epenthetic [i]

Source words	Loanwords
*match [*mætʃ]	[mattɕ*i]*
*beach [bi*tʃ]	[bi:tɕ*i]*
*edge [*ɛdʒ]	[eddʑ*i]*
*page [*peɪdʒ]	*[pe:*dʑ*i]*

The palato-alveolar affricates, [*tʃ*] and [*dʒ*], in the source words are adapted as [tɕ] and [dʑ] in Japanese loanwords, both of which are palatal affricates. I use the constraint of IDENT-IO for this featural alteration from [*tʃ*] and [*dʒ*] to [tɕ] and [dʑ]. Preceded by these affricates, [i] is physically more economical to epenthesize than [ɯ] because the articulation of palatalized segments raises the front of the tongue toward a position closer to that for [i] on a primary gesture. (Y. Kobayashi, 2005; Kubozono, 1999; Ladefoged & Maddieson, 1996; Lovins, 1975). This easier movement of the tongue draws the constraint, PALATAL-FRONT: vowels after palatals should be front. I add these two constraints, IDENT-IO and PALATAL-FRONT above, to the constraint-ranking in Tableau 5. The two constraints, IDENT-IO and PALATAL-FRONT should be higher ranked than all the constraints in Tableau 5 so that the generation of [ɯ] is canceled.

Table 8.
*OT Tableau 6. Generation of [i] in Japanese: PALATAL-FRONT, IDENT-IO >> *ROUND, *LOW >> *FRONT >> HIGH, BACK*

…tʃ/dʒ	PALATAL-FRONT	IDENT-IO	*ROUND	*LOW	*FRONT	HIGH	BACK
tɕ/dʑ a	*!	*		*		*	*
☞ tɕ/dʑ i		*			*		*
tɕ/dʑ ɯ	*!	*					*
tɕ/dʑ e		*			*	*!	*
tɕ/dʑ o	*!	*	*			*	

Context-dependent Epenthetic Vowel [o]

Another context-dependent epenthetic vowel in Japanese loanwords is [o], which is epenthesized after alveolar stops, [t] and [d]. Examples are shown below.

Table 9.
Loanwords with Epenthetic [o]

Source words	Loanwords
eight [*eɪt]*	[eito]
*mattress [*mætrɪs]	[mattoresɯ]
old [*oʊld]*	[o:ɾɯ*do]*
*trend [*trɛnd*]*	[to*ɾendo]*

If the better-qualified epenthetic vowel , [ɯ] or [i], were epenthesized after [t] or [d], the output would be [tɯ, dɯ] or [ti, di], which the Japanese syllable inventory does not have. As mentioned earlier, the constraint, IS, only allows for the segment that exists in the inventory of the language in question. In this study, IS could be expanded to SYLLABLE INVENTORY STRUCTURE (SIS), that is, using the syllables that exist in the syllable inventory of the language in question. SIS should be more suitable for the analysis of a syllable-oriented (not segment-oriented) language such as Japanese. The possible outputs, [tɯ, dɯ, ti, di], would violate SIS. Also, if [ɯ] or [i] were still epenthesized after [t] or [d] without violating SIS, [t] and [d] would have to be altered. As a result, the outputs would be [tsɯ, dzɯ, tɕi, dʑi], which indeed exist in the Japanese syllable inventory, but the alteration of [t] or [d] violates IDENT-IO. In order not to violate SIS and IDENT-IO, Japanese speakers avoid epenthesizing [ɯ] or [i]. Instead, they epenthesize another non-low vowel, [o], which has the third shortest intrinsic duration after [ɯ] and [i] as Han's scale shows in (1).

We add the constraints, SIS, to the ranking in Tableau 6. This constraint is ranked as high as PALATAL-FRONT and IDENT-IO. As shown in the tableau below, they evaluate the candidates with [t] and [d] that precede the epenthetic site. However, we encounter a problem with this tableau; this constraint-ranking generates [e], not [o].

Table 10.
*OT Tableau 7. Failure to Generate [o] in Japanese: SIS, PALATAL-FRONT, IDENT-IO >> *ROUND, *LOW >> *FRONT >> HIGH, BACK*

...t/d	SIS	PALATAL-FRONT	IDENT-IO	*ROUND	*LOW	*FRONT	HIGH	BACK
t/d a					*!		*	*
t/d i	*!					*		*
t/d ɯ	*!							*
💣 *t/d* e						*	*	*
√ *t/d* o				*!			*	
tɕ/dʑ i			*!			*		*
ts/dz ɯ			*!					*

Note. 💣 stands for an unpredicted optimal candidate, which nevertheless is generated. √ stands for a predicted optimal candidate, which nevertheless is not generated.

Here, we reexamine the quality of the Japanese [o]. We know the Japanese [o] is more centralized than cardinal [o]. Ladefoged and Maddieson (1996) state that, although back vowels are usually rounded, sometimes a language has relaxed the linkage between backness and rounding.3 In Whitman's (1985) study on the pre-old Japanese sound system, [o] is located at the center of the pre-old Japanese vowel inventory, like schwa. Also, Hamano (1998) finds that the Japanese [o] does not accompany protrusion or tenseness. Referring to these observations, I assume that the roundedness of the Japanese [o] is too weak to violate *ROUND. In addition, we find that HIGH needs to be placed higher than BACK to generate right vowels. This reexamination of [o] will rewrite the Tableaux 5, 6 and 7 as below. They generate the correct optimal forms.

Table 11.
*OT Tableau 8. (OT Tableau 5 with modification) Generation of [ɯ] in Japanese: *ROUND, *LOW >> *FRONT >> HIGH >> BACK*

CVC	*ROUND	*LOW	*FRONT	HIGH	BACK
CVCa		*!		*	*
CVCi			*!		*
☞ CVCɯ					*
CVCe			*!	*	*
CVCo				*!	

Table 12.
*OT Tableau 9. (OT Tableau 6 with modification) Generation of [i] in Japanese: PALATAL-FRONT, IDENT-IO >> *ROUND, *LOW >> *FRONT >> HIGH >> BACK*

...*tʃ/dʒ*	PALATAL-FRONT	IDENT-IO	*ROUND	*LOW	*FRONT	HIGH	BACK
tɕ/dʑ a	*!	*		*		*	*
☞ tɕ/dʑ i		*			*		*
tɕ/dʑ ɯ	*!	*					*
tɕ/dʑ e		*			*	*!	*
tɕ/dʑ o	*!	*				*	

Table 13.
*OT Tableau 10. (OT Tableau 7 with modification) Generation of [o] in Japanese: SIS, PALATAL-FRONT, IDENT-IO >> *ROUND, *LOW >> *FRONT >> HIGH >> BACK*

...*t/d*	SIS	PALATAL-FRONT	IDENT-IO	*ROUND	*LOW	*FRONT	HIGH	BACK
t/d a					*!		*	*
t/d i	*!					*		*
t/d ɯ	*!							*
t/d e						*!	*	*
☞ *t/d* o							*	
tɕ/dʑ i			*!			*		*
ts/dz ɯ			*!					*

3 Ladefoged and Maddieson (1996) refer to Japanese [ɯ] as this back unrounded vowel. This paper considers [ɯ] as a central vowel.

Overall, this set of constraints in the ranking successfully generates the epenthesis of [ɯ] as the default vowel, [i] after palatals, and [o] after alveolar stops.

In the analysis above, we constructed the ranking of SIS, PALATAL-FRONT, IDENT-IO >> *ROUND, *LOW >> *FRONT >> HIGH >> BACK. To this ranking, we can add the four general constraints for epenthesis, *COMPLEX, CODACOND, MAX-IO and DEP-IO. This completes the single ranking as shown in Table 14.

Table 14.
Constrains and Unified Ranking for Epentheses of [ɯ], [i] and [o] in Japanese

Constraints	Functions
*COMPLEX, CODACOND	(←Repair coda / consonant clusters)
>> MAX-IO	(←Do not delete, do epenthesis)
>> SIS, PALATAL-FRONT, IDENT-IO	(←Epenthesize context-dependent vowels)
>> *ROUND, *LOW >> *FRONT >> HIGH >> BACK	(←Epenthesize context-free default vowel)
>> DEP-IO	

This constraint-ranking should explain how Japanese speakers modify English phonotactics, which do not fit the Japanese phonological system.

Irregular Adaptations
Epenthesis of [i] after Dorsal Consonant

In this section, I will discuss some unstable, shaky and irregular patterns of vowel epenthesis in Japanese loanwords. One of the irregular patterns is [i]-epenthesis after non-palatals. This can be explained by historical change of vowel epenthesis. Example loanwords with the irregular [i]-epenthesis are shown below.

Table 15.
Loanwords with Irregular Epenthesis of [i]

Source words	Loanwords
cake [keɪk]	[ke:ki]
steak [*steɪk*]	[sɯte:ki]
*brake /*breɪk]	[bɯ*re:ki]*
ink [ɪŋk]	*/in*ki]
shake [ʃeɪk] (as a fast food)	*/*se:ki]

When observing the loanwords and source words above, we see that [i] is epenthesized after the voiceless dorsal consonant [k], and the neighboring vowels of [k] in the source words are front vowels such as [ɪ] and [ɛ]. This appears to be a vowel harmony of the back feature.

In old Japanese, there used to be vowel harmony in the native vocabulary. Ichikawa observes that, "[i] was added where the neighboring vowel was a front vowel like [e] or [i], and . . . [ɯ] or [o] occurred when it was preceded or followed by a back vowel. In this we see a sort of vowel harmony" (1930, p. 183). Thus, this epenthesis of [i] in some loanwords such as the ones in Table 15 is considered to be the residue of the vowel harmony in old Japanese.4 It is considered that the use of the epenthetic vowels from vowel harmony have been diachronically replaced by the use of the default epenthetic vowel. Some doublet loanwords and homophone loanwords evidence this analysis.

Table 16.
Doublets and Homophones

1. Doublets

Source words	Old loanwords (Vowel harmony)		Recently-made doublet (Default [ɯ])
shake [ʃeɪk] (as a fast food)	*[*se:ki]	→	*[*ʃeikɯ]
ink [ɪŋk]	*[in*ki]	→	[inkɯ]

2. Homophone

Source words	Old loanword (Vowel harmony)	Recently-made loanword (Default [ɯ])
brake *[*breɪk]	[bɯ*re:ki]*	
break [breɪk]		[bɯreikɯ]

The loanwords to the right in the tables above were more recently coined than the ones to the left. The recent loanwords exhibit the default epenthetic vowel [ɯ] while the old loanwords exhibit vowel harmony although the pronunciations of the source words are the same. Based on this irregular [i]-epentheis after non-palatals, I introduce a constraint, HARMONY, which used to be higher ranked but has been lowered.

Table 17.
*OT Tableau 11. Old Loanwords: Output [ki] as [bu*re:ki*] (< break [breɪk])*

…ik	SIS	PALATAL -FRONT	IDENT -IO	HARMONY	*ROUND	*LOW	*FRONT	HIGH	BACK
…ik a				*!		*		*	
☞…ik i							*		*
…ik ɯ				*!					*
…ik e				*!			*	*	*
---ik o				*!				*	

4 Japanese also has some loanwords that show vowel harmony other than [i]. An example is [sarada] (< "salad" [sæləd]).

Table 18.
OT Tableau 12. Recent Loanwords: Output [kɯ] as [bureikɯ] (< break [breɪk])

…ik	SIS	PALATAL -FRONT	IDENT -IO	*ROUND	*LOW	*FRONT	HIGH	BACK	HARMONY
…ik a					*!		*		*
…ik i						*!		*	
☞…ik ɯ								*	*
…ik e						*!	*	*	*
---ik o							*!		*

Certain questions remain. If vowel harmony was diachronically replaced by the use of the default epenthetic vowel, why does epenthetic [i] still survive in some loanwords, particularly in the context with [k]? First, we know that [i] is highly qualified as an epenthesized vowel, generally speaking: [i] is [+high, -round], and the intrinsic duration of [i] is the second shortest next to [ɯ]. In addition, referring to Rose and Demuth's (2006) study on the vowel epenthesis in Sesotho, a dorsal consonant does not block harmony between the neighboring vowels. This "transparent" quality of the dorsal consonant and the high qualification of [i] as an epenthetic vowel might have let [i] stay in the epenthetic slots in many loanwords.

Adaptation of [r]

The adaptation of [r] shows various patterns. In some loanwords, the default vowel [ɯ] is epenthesized after [r], as shown below in Table 19-1. However, in some other loanwords, [r] is deleted, and the preceding vowel is lengthened. The lengthened preceding vowel fills in the position for [r], as shown in Table 19-2. *Also, there are some other loanwords, in which [r] is replaced with [a] as in Table 19-3.*

Table 19.
Modifications of [r]

	Modifications		Source words	Loanwords
1.	*r* → ɾɯ / { V_σ, V_C }		allergy [ælərdʒi]	[areɾɯgi:]
			tornado [tɔrneɪdoʊ]	[toɾɯne:do]
2.	r → V1 / { V1_σ, V1_C }		guitar [*gɪtar]*	[gita:]
			soccer [*sɒkər]*	[sakka:]
			fork [*fɔrk]*	*[fo:k*ɯ]
3.	r → a / { V_σ, V_C }		store [*stɔr, stoʊr]*	[sɯtoa]
			core [kɔr, koʊr]	[koa]
			fair [fɛər]	[fea]

In Table 19-2 and 19-3, the adaptation of [r] does not utilize epenthetic vowels, and it becomes a vowel.5

5 It is common that English syllabic /r/ is analyzed as a rhotic vowel.

Experiment

An experiment was conducted in order to test whether the modifications that loanwords exhibit are applicable to Japanese speakers' pronunciation of English words. In particular, the experiment tested whether native Japanese speakers' modifications with epenthetic vowels can be predicted by the constraint-ranking in Table 14. Fifteen native Japanese speakers participated in the experiment. They were orthographically provided with 70 nonsense words spelled in alphabets, which included word-final codas and word-initial consonant clusters.6 Their task was to rewrite those non-words in Japanese characters. A Japanese character represents a mora-bearing unit, which is basically a syllable such as V or CV. In its syllabic writing system, there is no way to write coda consonants or consonant clusters, so the participants had to either delete consonants or epenthesize vowels when rewriting the stimuli. The nonsense words in the experiment were in ten different conditions as shown below. Also, the predicted epenthesized vowels according to the constraint-ranking in Table 14 are shown.

Condition (iv) might show some epenthetic [i] due to the vowel harmony, but this constraint should be ranked low in modern pronunciation, so default [ɯ]-epenthesis should be the most cases.

6 The ages of the participants ranged from 19 to 42. They were the students of the undergraduate program and ESL school, which belong to the University of South Carolina and Clemson University.

Table 20.
Conditions for Experiment

a. Word-final Coda

Stimuli (Nonsense words)	Predicted epenthetic vowels
(i) [b, f, m, p, q, s, v, z] (e.g. *gamb, ktnof*)	[ɯ] (default vowel)
(ii) [*tʃ*, *dʒ*], which were spelled with *ch* and *dge*, respectively. (e.g. *consuch, zodge*)	[i] (after palato-alveolar affricates: PALATAL-FRONT, IDENT-IO)
(iii) [t, d] (e.g. *dmlt, zod*)	[o] (after alveolar stops: SIS, IDENT-IO)
(iv) [k, g] (dorsal) preceded by *i* (e.g. *ponkik, pog*)	[ɯ] (default vowel)
(v) [r] (e.g. *hmor*)	[ɯ] or alternation of [r] to vowels (either preceding vowel or [a])

b. Word-initial Consonant Cluster

First consonants in stimuli (Nonsense words)	Predicted epenthetic vowels
(i) [b, f, m, p, q, s, v, z] (e.g. *bkautu, fmil*)	[ɯ] (default vowel)
(ii) [*tʃ*, *dʒ*], which were spelled with *ch* and *j*, respectively. (e.g. *chki, jktap*)	[i] (after palato-alveolar affricates: PALATAL-FRONT, IDENT-IO)
(iii) [t, d] (e.g. *tnoman, dvolt*)	[o] (after alveolar stops: SIS, IDENT-IO)
(iv) [k, g] (dorsal) followed by *i* (e.g. *ktmos, gsamol*)	[ɯ] (default vowel)
(v) [r] (e.g. *rbran*)	[ɯ] or alternation of [r] to vowels (either preceding vowel or [a])

Results of Word-final Codas

The table below shows the types of the vowels that the participants epenthesized.

Table 21.
Types of Epenthesized Vowels: Word-final Coda

	Epenthesized vowels					Other types of repair			
	ɯ	i	o	a	e	r→a	r→ :	delete	n.a.
(i) [b, f, m, p, q, s, v, z]	**96.7%**					-	-	0.8%	2.5%
(ii) [*tʃ*] or [*dʒ*] spelled with *ch* and *dg*	12.2%	**85.6%**				-	-		2.2%
(iii) [t] or [d]	5.6%		**91.1%**			-	-	2.2%	1.1%
(iv) [k] or [g] (Dorsal) preceded by *i*	**95.6%**	3.3%				-	-		1.1%
(v) [r]	**28.9%**			2.2%		**33.3%**	**35.6%**		

Note. Expected epentheses are in bold. Unexpected epentheses are underlined.

Despite the small number of discrepancies, the result mostly agreed with my constraints in Table 14. For the condition (ii) and (iii) with affricates and alveolar stops, some participants epenthesized the default [ɯ] instead of [i] and [o], respectively. This indicates that the participants place the constraints for the default vowel higher than PALATAL-FRONT, IDENT-IO, and SIS in their rankings.

Results of Word-initial Consonant Clusters

The adaptations of the word-initial consonant clusters were not as simple as coda. The results are shown below.

Table 22.
Types of Epenthesized Vowels: Word-initial Consonant Cluster[7]

	Epenthesized vowels					Other types of repair			
	ɯ	i	o	a	e	r→a	r→ :	delete	n.a.
(i) [b, f, m, p, q, s, v, z]	**90.8%**	1.7%	5.0%	1.7%	0.8%	-	-		
(ii) [*tʃ*] or [*dʒ*] spelled with *ch* and *j*	23.7%	**34.4%**	4.4%	6.3%	3.0%	-	-	1.5%	26.7%
(iii) [t] or [d]	32.2%	7.8%	**45.6%**		3.3%	-	-	8.9%	2.2%
(iv) [k] or [g] (Dorsal) preceded by *i*	**71.0%**	**15.6%**		2.2%		-	-	10%	
(v) [r]	**71.1%**		6.7%	6.7%		**8.9%**	-	6.7%	

Note. Expected epentheses are in bold. Unexpected epentheses are underlined.

The results of the repairs for consonant clusters exhibit a great variety of epenthetic vowels. The [ɯ]-epenthesis after affricates ([*tʃ, dʒ*]) and alveolar stops ([t, d]) can be explained. The constraints that generate the default [ɯ] were higher ranked than SIS, IDENT-IO and PALATAL-FRONT in many participants. However, how can we explain the epenthesis of [e], [a], and other unexpected vowels in the word-initial consonant cluster condition?

A closer observation of the unexpected epenthetic vowels in the results revealed the effect of vowel harmony: 58.9% of the unexpected epenthetic vowels exhibited vowel harmony effect as summarized below.

7 For the condition (ii), only 73.3% of the answers interpreted the spelling of /ch/ and /j/ as [*tʃ*] and [*dʒ*]. Other answers interpreted /ch/ as [k] or [*ʃ*].

Table 23.
Epentheses with Vowel Harmony: Word-initial Consonant Cluster

	Total Vowel Harmony	Partial (back feature) Vowel Harmony	No Harmony
(i) [b, f, m, p, q, s, v, z]	100%		
(ii) [*tʃ*, *dʒ*] spelled with *ch* and *j*	71.1%		28.9%
(iii) [t, d]		27.8%	72.2%
(v) [r]	44.4%		55.6%

This indicates that in the present day Japanese is still influenced by vowel harmony to some extent, although the constraint HARMONY has been lowered as shown in the results of condition (iv) in Tables 21 and 22. However, the question is why vowel harmony was much more present for word-initial consonant clusters than for word-final codas. In observing this, we should think about the directionality of vowel harmony. The current experiment shows the regressive harmony (in the cluster condition) much more often than the progressive harmony (in the coda condition). This could be explained by Hansson (2001), who states that regressive (right-to-left) assimilation is the default direction of harmony. Further, the constraint, HARMONY, might have to be divided into REGRESSIVE HARMONY and PROGRESSIVE HARMONY and be ranked independently as REGRESSIVE HARMONY >> PROGRESSIVE HARMONY. This can be a topic of a further study.

CONCLUSION

This paper showed the patterns of native Japanese speakers' modifications of the pronunciation of English words based on the analysis of loanwords. Japanese speakers utilize the context-free default epenthetic vowel [ɯ] and the context-dependent vowels [i] and [o]. The single-constraint-ranking that explains these epentheses that was shown in Table 14 is repeated: *COMPLEX, CODACOND >> MAX-IO >> SIS, PALATAL-FRONT, IDENT-IO >> *ROUND, *LOW >> *FRONT >> HIGH >> BACK >> DEP-IO.8 In addition, the experiment in this study finds that, to some extent, vowel harmony is still active in native Japanese speakers' selections of epenthetic vowels, especially in pronouncing consonant clusters in English. This implies that a constraint of vowel harmony could be included in the constraint-ranking above. Further study about this harmony effects in Japanese speakers' pronunciation is warranted.

ACKNOWLEDGEMENTS

8 This constraint-ranking explains the patterns of vowel epenthesis. This does not explain the alteration of [r] to vowels.

Many thanks to Professor Eric Holt for supporting this research. I would like to express my gratitude to the editors and the reviewer for their patience.

ABOUT THE AUTHOR

Shinichi Shoji was a Japanese adjunct instructor from 2003 to 2006 and was a lecturer from 2006 to 2010 at University of North Carolina at Charlotte. He began work on his Ph. D in linguistics in 2010 at University of South Carolina with the purpose of merging his interests in formal linguistics and language teaching. Contact information: 901 Sumter St. Byrnes 207, Columbia, SC 29208. (704) 975-3191. shoji@email.sc.edu

REFERENCES

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 206-244). Oxford: Blackwell.

Carr, P. (1999). *English phonetics and phonology: An introduction*. Malden, MA: Blackwell.

Hamano, S. (1998). *The sound-symbolic system of Japanese*. Stanford, CA: Center for the Study of Language and Information (CSLI) publications.

Han, M. S. (1962). The feature of duration in Japanese. *The Study of Sounds, 10*, 65-80.

Hansson, G. Ó. (2001). *Theoretical and typological issues in consonant harmony*. Unpublished doctoral dissertation, UC Berkeley.

Harrison, K. D. and A. Kaun. (2000). Pattern-responsive lexicon optimization. *North Eastern Linguistic Society, 30,* 327-340.

Hirayama, M. (2003). Contrast in Japanese vowels. *Toronto Working Papers in Linguistics 20*. Retrieved from http://twpl.library.utoronto.ca/index.php/twpl/article/view/6234

Ichikawa, S. (1930). The pronunciation of English loan-words in Japanese. In N. Bøgholm, A. Brusendorff & C. A. Bodelsen (Eds.), *A grammatical miscellany offered to Otto Jespersen on his seventieth birthday* (pp. 179-190). Copenhagen: Levin & Munksgaad.

Kager, R. (1999). *Optimality theory*. Cambridge: Cambridge University Press.

Kang, Y. (2011). Loanword phonology. In M. Ostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology (chap. 95)*. Retrieved from http://www.companiontophonology.com/subscriber/uid=951/tocnode?id=g9781405184236_chunk_g978140518423697

Kenstowicz, M. (2003). Salience and Similarity in Loanword Adaptation: A Case Study from Fijian. Retrieved from http://roa.rutgers.edu/files/609-0803/609-KENSTOWICZ-0-0.PDF

Kobayashi, Y. (2005). *Nichiei gairaigo no hatsuon* [*Pronunciation of Japanese loanwords from English*]. Hiroshima, Japan: Keisuisha.

Kubozono, H. (1999). *Nihongo Onseigaku* [*Japanese phonetics*]. Tokyo, Japan: Iwanami Shoten.

Ladefoged, P., & I. Maddieson. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge: MIT Press.

Lombardi, L. (2002). Markedness and the typology of epenthetic vowels. Retrieved from http://roa.rutgers.edu/files/578-0203/578-0203-LOMBARDI-0-1.PDF

Lovins, J. B. (1975). *Loanwords and the phonological structure of Japanese*. Bloomington: Indiana University Linguistic Club.

Mori, M. (1929). *The pronunciation of Japanese*. Tokyo, Japan: The Harald-sha.

Paradis, C. & D. LaCharité. (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics, 33*, 379–430.

Rose, Y. & K. Demuth. (2006). Vowel epenthesis in loanword adaptation: representational and phonetic considerations. *Lingua 116*, 1112-1139.

Sapir, E. (1921). *Language: An introduction to the study of speech*. Rahway, NJ: Quinn & Boden Company.

Steriade, D. (2009). *The phonology of perceptibility effects: the P-map and its consequences for constraint organization*. In K. Hanson & S. Inkelas (Eds.), The nature of the world: studies in honor of Paul Kiparsky (pp. 151-180). Cambridge, MA: MIT Press.

Vance, T. J. (2008). *The sounds of Japanese*. Cambridge, UK: Cambridge University Press.

Whitman, J. B. (1985). *The phonological basis for the comparison of Japanese and Korean*. Unpublished doctoral dissertation, Harvard University.

Yoshida, Y. (2006). Accents in Tokyo and Kyoto Japanese vowel quality in terms of duration and licensing potency. *SOAS Working Papers in Linguistics 14,* 249-264.

Keyworth, P. (2014). The acoustic correlates of stress-shifting suffixes in native and nonnative English: Some preliminary findings. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 104-120). Ames, IA: Iowa State University.

THE ACOUSTIC CORRELATES OF STRESS-SHIFTING SUFFIXES IN NATIVE AND NONNATIVE ENGLISH: SOME PRELIMINARY FINDINGS

Paul R. Keyworth, Saint Cloud State University

Although laboratory phonology techniques have been widely used to discover the interplay between the acoustic correlates of English Lexical Stress (ELS) – fundamental frequency, duration, and intensity - studies on ELS in polysyllabic words are rare, and cross-linguistic acoustic studies in this area are even rarer. Consequently, the effects of language experience on L2 lexical stress acquisition are not clear. This investigation of adult Arabic (Saudi Arabian) and Mandarin (Mainland Chinese) speakers analyzes their ELS production in tokens with seven different stress-shifting suffixes. Stress productions are then systematically analyzed and compared with those of speakers of Midwest American English using the acoustic phonetic software, *Praat*. In total, one hundred subjects participated in the study, spread evenly across the three language groups. Nonnative speakers completed a sociometric survey prior to recording so that statistical sampling techniques could be used to evaluate acquisition of accurate ELS production. The speech samples of native speakers were analyzed to provide norm values for cross-reference and to provide insights into the relative salience hierarchy of the three acoustic correlates of stress. The results support the notion that differences in lexical stress production in varieties of L2 English can be directly attributed to differences in the L1 sound system; hence, nonnative ratios of the acoustic cues lead to accented speech. Furthermore, the findings suggest that native-like command of ELS can be acquired by proficient L2 learners via increased L2 input.

INTRODUCTION

Certain suffixes in English cause a shift in stress in the root morpheme to the syllable directly preceding the suffix, and have hence been labeled *stress-shifting suffixes* by pronunciation experts (e.g., Celce-Murcia, Brinton, and Goodwin, 1996, Kreidler, 2004). They have claimed that the resultant shift in stress in turn causes a change in the neutralization or vowel reduction in the unstressed syllable. However, these claims about lexical stress shifts have not yet been supported quantitatively by laboratory phonology. Furthermore, quantification of stress shifts in suffixal derivations has scarcely been utilized as a tool in the acoustic phonetic characterization of accentedness in the speech of second language (L2) English speakers from different first language (L1) backgrounds.

Background and Need for the Study

The acoustic correlates of lexical stress are fundamental frequency (hereinafter F0), duration, and intensity(9). Although various laboratory phonology studies have analyzed these cues in the

(1) According to Ladefoged (2006), for all intents and purposes, F_0 is synonymous with pitch (measured in Hertz, Hz), duration means vowel length (measured in milliseconds, ms), and intensity equates to loudness (measured in Decibels, db). Frequency is the number of cycles of variations in air pressure in one second, and pitch is the auditory feature that allows listeners to

context of English Lexical Stress (ELS) in general, they have not explored the acoustic properties of the full range of stress-shifting suffixes in the lexicon. In fact, studies on ELS in *polysyllabic* words in general have largely been ignored in favor of disyllabic minimal stress pairs, as in Fry's seminal works (1955, 1958).

Moreover, there are few studies comparing the productions of stress-shifting suffixes by native speakers of English (NSE) and nonnative speakers of English (NNSE), and the effects of L1 background and amount of L2 exposure/input on pronunciation accuracy have not been fully explored. Munro and Derwing (1995) have emphasized the need for further studies on the features of L2 pronunciation that have the most significant effect on intelligibility in English. Similarly, Ramus, Nespor, and Mehler (1999) have reiterated the demand for a more determined approach from acoustic phoneticians in order to ascertain the properties of stress in different languages.

Although the extent of L2 accentedness is related to many determinants, including language environment and age of speakers, the main mediator of individual differences in L2 accents is the "sound system" of their L1 (Zhang, Nissen, & Francis, 2008, p. 4498). For example, there is growing evidence to suggest that Mandarin L1 speakers have problems pronouncing L2 English stress contrasts because of "strong interference from the Mandarin *tonal* system" (Zhang et al., 2008, p. 4500). As Zhang et al. have stated, even when syllabic stress *is* placed appropriately by Mandarin NNESs, they have problems manipulating the acoustic correlates of stress in a native-like manner.

Conversely, various phonetic studies on rhythmic typology strongly indicate that Arabic is a stress-timed language that is "a very likely language to exhibit the same correlates to stress as does English" (de Jong & Zawaydeh, 1999, p. 5). For these reasons, it is of interest to investigate whether Arabic L1 NNSE typically produce the acoustic correlates of stress more accurately than Mandarin L1 NNSE. Notwithstanding, Shemshadsara (2011) has provided statistical empirical evidence that Iranian Persian speakers have more difficulty placing stress in words with stress-shifting suffixes than words with neutral-suffixes (Burzio, 1994). Thus, these data further necessitate the need for a thorough acoustic analysis of stress-shifting suffixes.

Purpose of the Study

First, this project aims to quantify lexical stress shifts and provide insight into the relative salience hierarchy of the acoustic correlates of ELS. To do this, vocalic F_0, duration, and intensity productions by native speakers of Midwestern American English (MWAE) dialect are analyzed in words containing stress-shifting suffixes.

From a second language acquisition (SLA) research perspective, the other purpose of this study is to identify problematic acoustic correlates for NNSE and observe whether there is a correlation between exposure to the L2, amount of L2 input, and/or L1 background and production accuracy of words containing stress-shifting suffixes. As Zhang et al. (2008) have succinctly noted, most

perceive a sound on a low-high spectrum where a sound with high frequency is realized as high pitched (Ladefoged, 2006).

research in the area of ELS "confound the phonological issue of stress placement with the phonetic problem of native-like stress production" (p. 4498). Thus, production accuracy here implies a twofold distinction:

1) L2 knowledge of where to place the stress in derived words
2) Native-like production of the acoustic correlates of stress.

Specifically, this study examines the acoustic correlates of ELS in productions of English words containing stress-shifting suffixes by Arabic L1 and Mandarin L1 NNSE.

Research questions

Q1. Do English suffixes such as <ious>, <ial>, <ian>, <ic>, <ical>, <ity>, and <ify> cause a shift in stress when spoken by native speakers of MWAE, and can this be observed quantitatively using acoustic measurements of F_0, duration, and intensity?

Q2. Which acoustic correlates are problematic for Arabic L1 and Mandarin L1 speakers when producing lexical stress contrasts in words containing stress-shifting suffixes?

Q3. Is there a correlation between amount of L2 exposure (years of residence in L2 environment) and/or amount of L2 input (years of L2 study) and *accurate production* of the three acoustic cues? I.e., do these variables show a large effect in:

a) Accurate placement of lexical stress in stress-shifting suffixal derivations.
b) Native-like production of the acoustic correlates in stress-shifting suffixed words.

METHOD
Participants

One hundred participants including 29 MWAE NSE (male n = 15; female n = 14), 38 Arabic L1 NNSE (male n = 15; female n = 23) from the Kingdom of Saudi Arabia (KSA), and 33 Mandarin L1 NNSE (male n = 15; female n = 18) from the People's Republic of China (PRC) participated in the experiment. Subjects were recruited primarily from the student population of the Intensive English Program (IEP), the College ESL program, and the regular undergraduate and graduate student body of a large public university in Minnesota. This study only used NNSE who grew up speaking MWAE dialect. Prior to acoustic data elicitation, each participant completed a short "factual" (Dörnyei, 2003, p.8) sociometric questionnaire. Figure 1 shows the total number of subjects by language and gender.

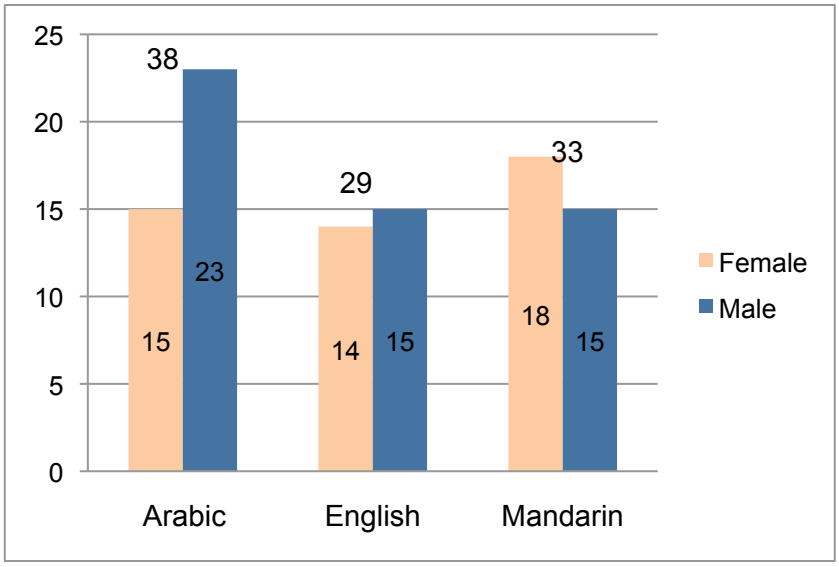


Figure 1. Participants by language and gender.

Materials
Stimuli

The procedure involved digitally recording (sampling rate of 44.1 kHz and quantization of 16 bits) productions of eight tokens, which included the stem word <HIStory> in addition to seven derived words formed by the addition of seven different stress-shifting suffixes: <hisTORic>, <hisTORical>, <histoRICity>, <hisTORian>, <hisTORify>, <hisTORial>, and <HisTORious>.

Participants produced the eight tokens randomly from the list in the carrier phrase *"Say* _____ *again"*. The carrier phrase was designed so that the words did not carry an onset rise or a pitch accent. According to Maeda (1976), an onset rise occurs at the start of a prosodic phrase, and a pitch accent denotes the main stress in the prosodic phrase. In this case, *"Say"* carries the onset rise, and *"again"* carries the pitch accent.

Procedure
Data collection

The researcher showed the eight tokens to the participants on eight separate cards. These cards were shuffled before each elicitation session to eliminate any potential ordering effects. Prior to recording, the participants did not receive any training in pronunciation of the words so that they would have to rely upon their own phonological knowledge of where to shift the primary stress.

Data Analysis

Acoustic phonetic analyses of the productions were performed in a similar methodology to Flege and Bohn (1989), Zhang et al. (2008), and Lee and Cho (2011) using *Praat* (Version 5.3.31). However, to the best of the author's knowledge, the proposed method of comparing vowels was novel in that the acoustic correlates in vowels with primary stress were examined in relation to those of *all* the other vowels in the utterance – as opposed to just one of the unstressed vowels

(e.g., Flege & Bohn, 1989; Lee & Cho, 2011). In other words, the [+tonic acc.] vowel in each token was acoustically compared to the [-tonic acc.] vowel (Ladefoged, 2001; See Appendix A). The rationale being, that if a vowel has primary stress, the acoustic cues should be prominent to all the other vowels in word.

Praat scripts were used to semi-automate delineation of vowels (Ryan, 2005) and retrieve the relevant stress analysis data (Yoon, 2008). First, the grid-maker script was run and vowel segments were defined in the series of spectrograms (Appendix B).

Once all the spectrograms had been annotated, Yoon's stress analysis script (2008) was run to collect the vocalic mean F0, mean intensity, and duration values. Thus, each individual production was measured for the mean F0/intensity/duration of the primary stressed vowel and the mean of the mean F0/intensity/duration values for all the other vowels (i.e., secondary stressed and unstressed).

Then, in order to answer *Research Questions 1* and *2*, the corresponding primary stressed to unstressed/secondary-stressed vowel ratios (i.e., [+tonic acc.] : [-tonic. acc] ratio) were calculated for each token to provide *vocalic relative stress values* for duration, intensity, and $F_{0.}$ This concurs with the methodologies of McClean and Tiffany (1973), Flege and Bohn (1989), and Lee and Cho (2011).

In order to answer *Research Question 2, and 3b,* the researcher followed the statistical methods employed by Zhang et al. (2008) and Lee and Cho (2011). To determine whether there was a statistical significance in the differences between the mean ratios of stressed to unstressed vowels for the independent variable of L1 background, the mean ratio of each acoustic factor was submitted to one-way ANOVAs. In addition, Tukey HSD tests were performed with a critical *p* value of 0.05.

At this point, it is important to note that not all of the data collected were included in the aforementioned calculations. If a token was mispronounced by a subject, it was categorized as either a *stress placement error* or a *pronunciation error* and entered into a separate data pool. The latter error type might also include stress placement errors but was deemed more serious with regards to intelligibility as it (also) included a deletion, and/or addition, and /or substitution of a segment or cluster of segments. Naturally, any productions containing errors had to be excluded from the main dataset; however, this data was useful to answer *Research Question 3a*. Pearson Product-Moment Correlations were used to determine the strength of relations and effect sizes (*r* values) for the operationalized variables of L2 exposure and L2 input.

RESULTS

Pronunciation and Stress Placement Errors

Figure 2 shows the percentage of errors for each language group. As one would expect, the NES made far fewer errors.

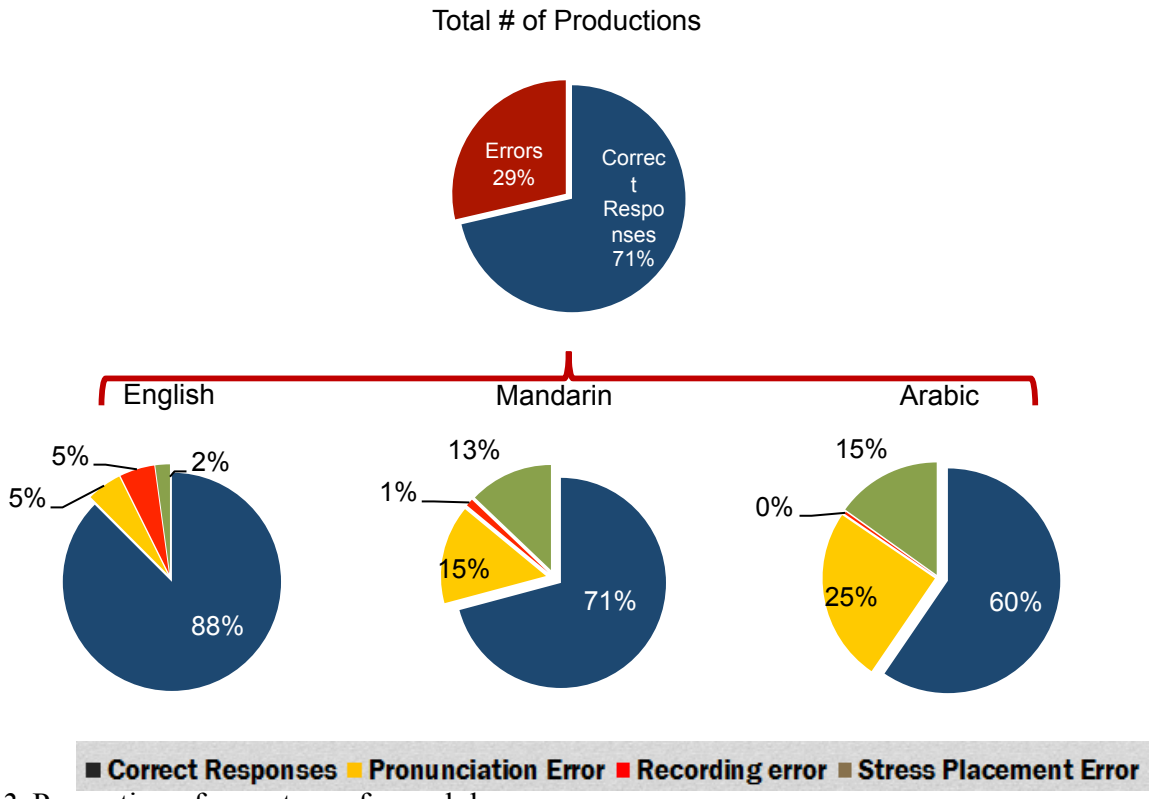


Figure 2. Proportion of error types for each language group.

Figure 3 shows the percentage of correct responses by token and language group. Also, as one might expect, <historicity> caused the most problems. In fact, NES were no better at accurately producing this word than Mandarin speakers. However, it is important to note that for the two nonsense words (i.e., <historious> and <historial>), the NES performed much better. It is the researcher's contention that although these are not real words, NES are able to use the stress-shifting rules that are stored in their lexicons.

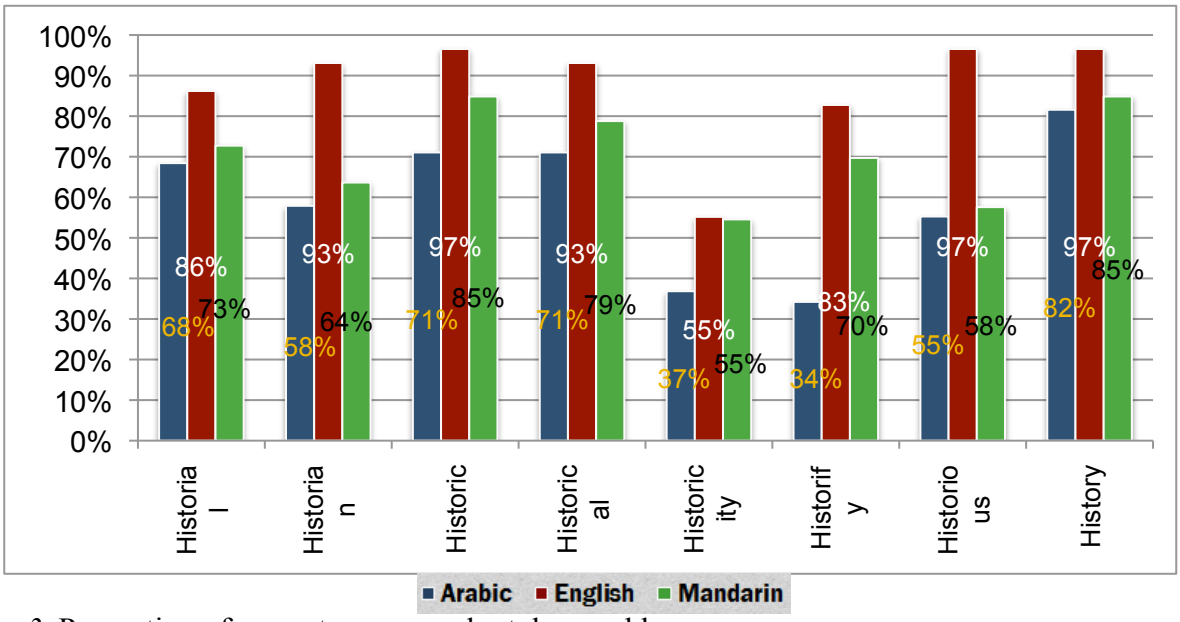


Figure 3. Proportion of correct responses by token and language group.

In Figure 4, the frequency of errors (pronunciation and stress placement) for all NNESs (Arabic and Mandarin L1 speakers) are plotted against years spent living in an English-speaking environment. No significant correlation was found (r= 0.09, *p* = n.s), perhaps because the study did not have a large enough range of participants with respect to this independent variable. Most participants had only spent between 0 and 1.5 years in an English-speaking country.

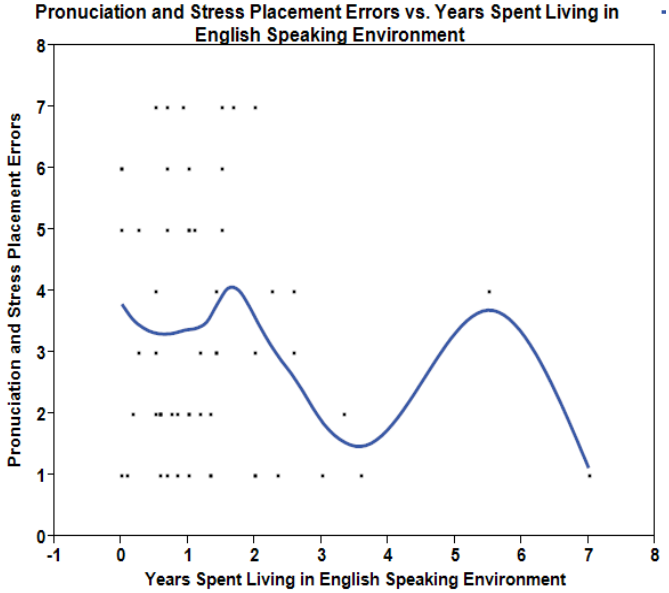


Figure 4. Frequency of pronunciation and stress-placement errors vs. years in L2 environment.

However, the smooth curve in Figure 5 shows that a significant correlation was found (r= - .26, p< .05.) for years of English language study and frequency of errors.

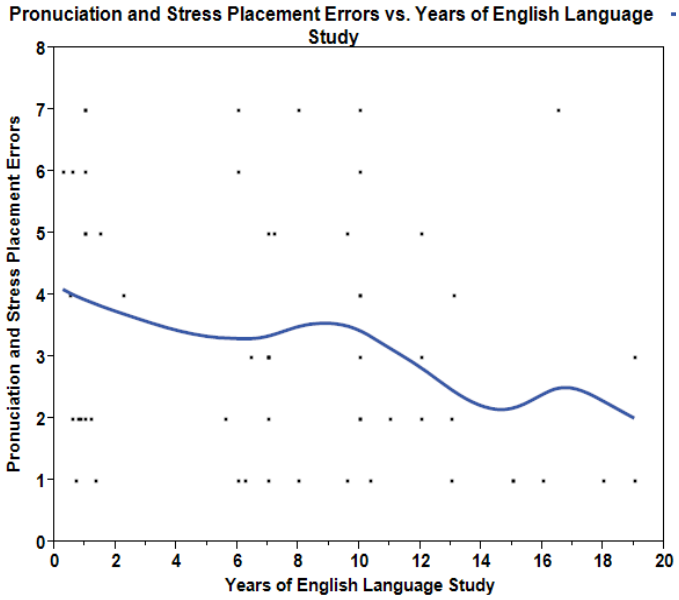


Figure 5. Frequency of pronunciation and stress-placement errors vs. years of L2 study.

Main Dataset

Table 1 shows the mean vocalic stress ratios for the three acoustic correlates of stress for all three language groups. On average, all 3 language groups produced primary-stressed vowels with higher F0, greater intensity, and longer duration. However, for each factor, the ratios were larger for English speakers than Mandarin and Arabic speakers. Thus, even though the NNESs were differentiating the stress in suffixed words, they did not do so to such a great extent as NES.

Table 1
Mean Ratios of Primary Stressed [+tonic acc.] to Non-primary Vowels [-tonic acc.] for the Three Acoustic Correlates by Language Groups

Language Group	F0	Intensity	Duration
Arabic	1.07	1.05	1.29
English	1.13	1.06	1.61
Mandarin	1.08	1.04	1.57

The mean ratios for each factor were entered into ANOVAs. The results are shown in Table 2.

Table 2
ANOVA Results for the Three Acoustic Correlates

Acoustic Correlate	df	F	Sig.
F0	484	2.799	.062
Intensity	484	5.246	.006

Duration	484	11.904	.000

One-way ANOVA/Tukey post-hoc tests found that neither Arabic nor Mandarin speakers were statistically different from native speakers with regards to F0 usage as an acoustic cue to ELS (Figure 6).

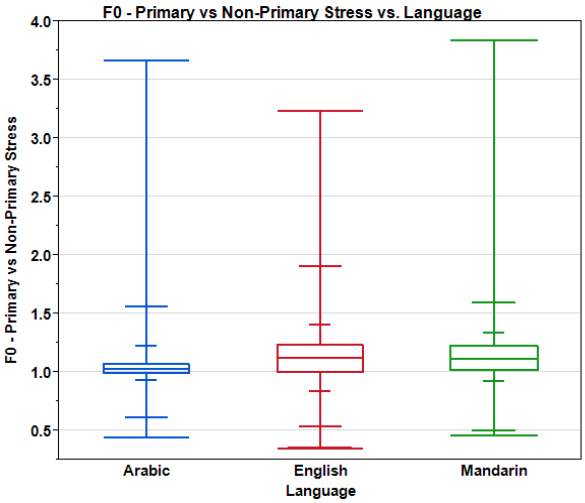


Figure 6. Comparative usage of F0 as a cue to ELS in stress-shifting suffixed words.

One-way ANOVA tests showed that there were significant differences between the language groups with regards to intensity as a prosodic cue, $F(2, 481) = 5.25$, $p < .01$. The Tukey HSD post hoc comparison revealed that Mandarin L1 NNES ($M = 1.04$, $SD = .05$) were statistically different from NES ($M = 1.06$, $SD = .04$) as they used a significantly lower ratio of intensity in stress contrasts (Figure 7).

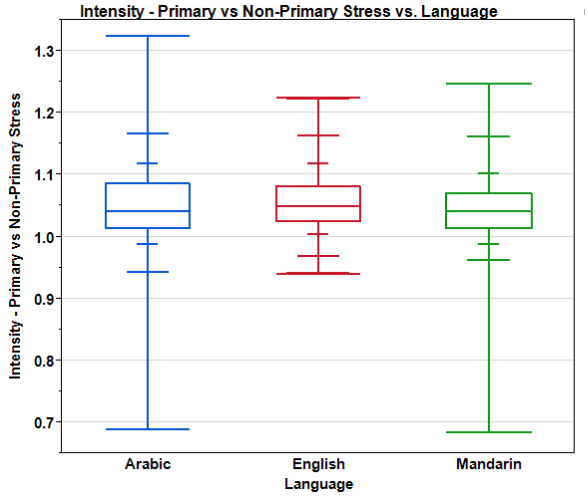


Figure 7. Comparative usage of intensity a cue to ELS in stress-shifting suffixed words.

Conversely, Figure 8 shows that it was Arabic L1 speakers who were statistically different from both Mandarin L1 speakers and NESs with respect to duration ratios of stressed to unstressed vowels. One-way ANOVA tests showed that there were significant differences between the language groups with regards to duration as a prosodic cue, $F(2, 481) = 11.90$, $p < .01$. The Tukey HSD post hoc comparison revealed that Arabic L1 NNES ($M = 1.29$, $SD = .70$) were statistically different from both NES ($M = 1.61$, $SD = .63$) and Mandarin L1 NNES

(*M* = 1.57, *SD* = .58) in their usage of durational stress contrasts (Figure 4); i.e., their relative vocalic stress ratios were much smaller.

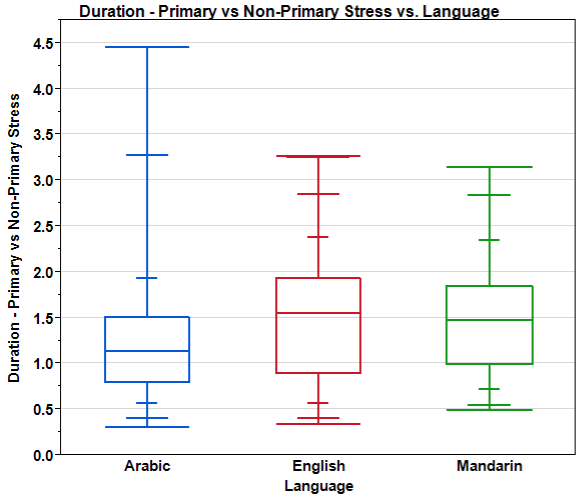


Figure 8. Comparative usage of duration as a cue to ELS in stress-shifting suffixed words.

Figures 9 and 10 show years of English L2 study versus native-like production of F0 (i.e., difference from the mean NES ratio) in lexical stress contrasts. While there was no significant correlation for Arabic L1 speakers, there was a significant correlation for Mandarin L1 speakers (r= - .49, *p*< .05). Therefore, the longer the Chinese subjects had claimed to have spent learning English, the more native-like they used pitch as an acoustic cue in lexical stress contrasts.

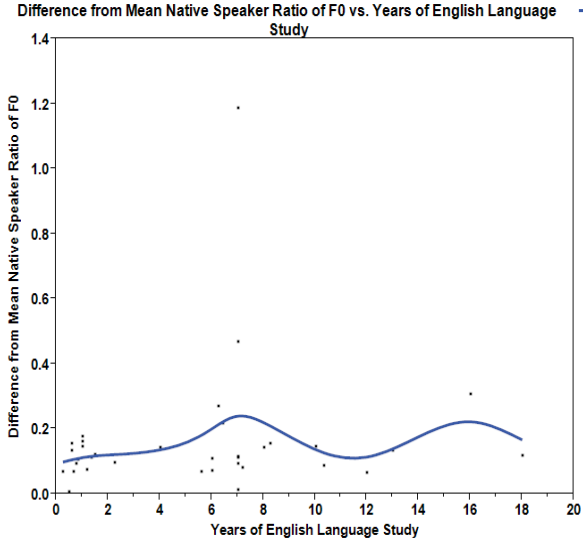


Figure 9. Difference of mean Arabic L1 speaker ratio of F0 from mean native speaker ratio of F0 vs. Years of L2 English study.

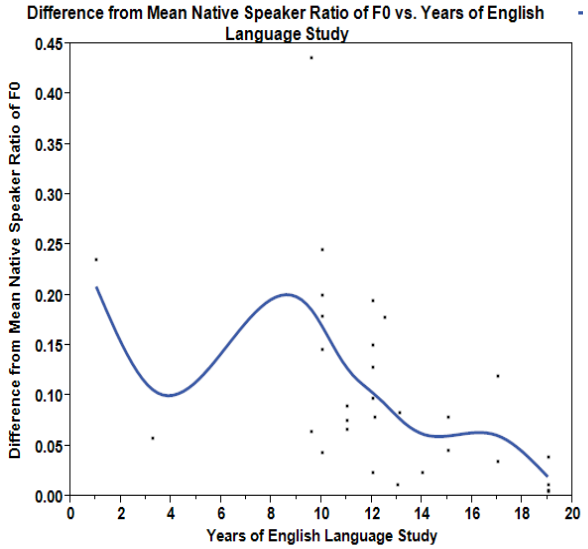


Figure 10. Difference of mean Mandarin L1 speaker ratio of F0 from mean native speaker ratio of F0 vs. Years of L2 English study.

Meanwhile, it was Arabic, not Mandarin L1 speakers, whose pronunciation improved with respect to duration usage as a result of increased English language study (Figures 11-12). Figure 11 shows that difference from mean NEs ratio of duration and years of English L2 study were significantly correlated for Arabic speakers, r= - .28, $p < .05$.

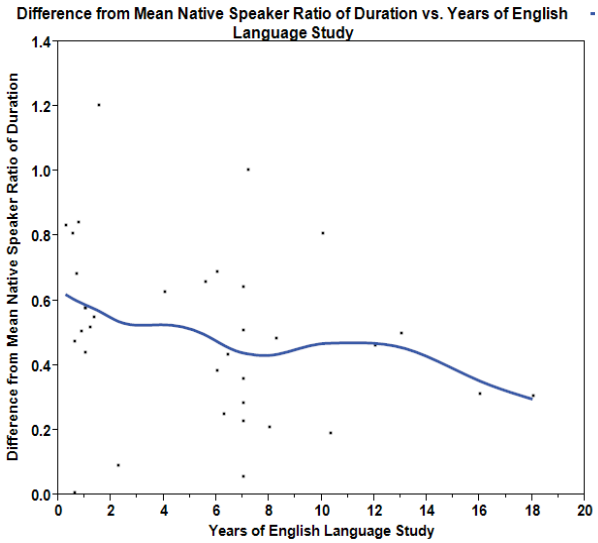


Figure 11. Difference of mean Arabic L1 speaker ratio of duration from mean native speaker ratio of duration vs. Years of L2 English study.

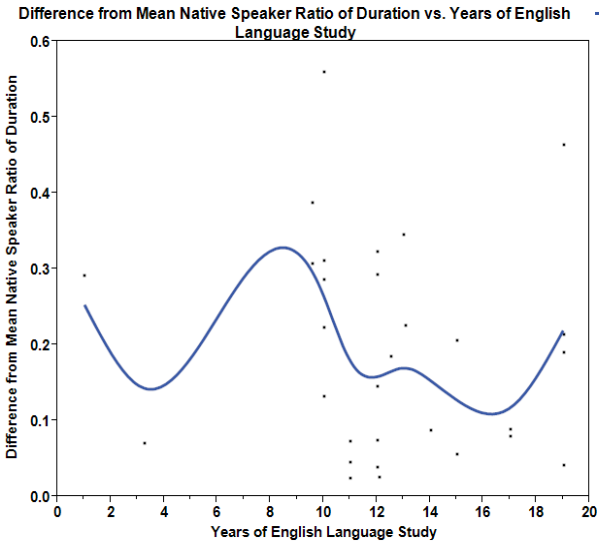


Figure 12. Difference of mean Mandarin L1 speaker ratio of duration from mean native speaker ratio of duration vs. Years of L2 English study.

DISCUSSION

With the study's research questions in mind, one can make the following observations.
Q1. Quantitative observation of lexical stress-shifts in derivations containing stress-shifting suffixes.

In accordance with the pronunciations stated in common textbooks and dictionaries, at least 90% of native speakers of MWAE placed the primary stress either on the vowel preceding the suffix in the derivations or on the first vowel in <history>. For these accurate productions, native speakers of MWAE on average produced primary-stressed vowels with higher F0, greater intensity, and longer duration than the average of the rest of the vowels combined. Thus, we can conclude that stress-shifts in words containing stress-shifting suffixes *can* be observed quantitatively.

Q2. Problematic acoustic correlates for Arabic L1 and Mandarin L1 speakers

On average, both Arabic L1 NNES and Mandarin L1 NNES produced primary-stressed vowels with higher F0, greater intensity, and longer duration, albeit with smaller vocalic relative stress ratios for each of the acoustic correlates than NES (Table 1). However, since the ratios for each acoustic cue are measured in different units (i.e., Hertz, decibels, and milliseconds) which are not calibrated to be perceptively equivalent or directly comparable, further statistical tests were necessary in order to determine problematic features of L2 speech. One-way ANOVAs revealed that Mandarin L1 speakers did not use intensity in a native-like manner while Arabic L1 speakers tended to *underuse* duration as an acoustic cue to ELS in stress-shifting suffixed words. The latter phenomenon is most likely caused by Saudi speakers not fully reducing vowels in stress-shifts which is a result of L1 transfer from the predictable stress-system of Arabic as suggested by several researchers (Zuraiq, 2005; Altmann, 2006; Bouchhioua, 2008).

Q3. Correlation between amount of L2 exposure and/or amount of L2 input and

a) Accurate placement of stress in stress-shifting suffixal derivations.

Figure 5 suggests that the longer learners of English have spent studying the language, the fewer pronunciation and stress-placement errors they make in stress-suffixed words. As mentioned, although amount of L2 exposure (i.e., years of residence in L2 country) did not yield statistical correlations, a study with a larger range of values for this variable may produce significant results.

b) Native-like production of the acoustic correlates of stress in stress-shifting suffixal derivations.

Mandarin L1 speakers used pitch more accurately (in a more native-like manner) as an acoustic cue to ELS, the longer they claimed to have spent learning English. Meanwhile, Arabic L1 speakers used duration more accurately (by reducing unstressed vowels), the longer they claimed to have spent studying English. Thus, it seems that through increased L2 acquisition, Saudi learners are able to overcome the detrimental effects of negative transfer.

CONCLUSION

This investigation supports the view that the relative salience hierarchy of the acoustic correlates of lexical stress is an important feature of English L2 accentedness. Clearly, the precise ratios of the three acoustic cues play an integral role in differentiating lexical stress patterns, and there does indeed appear to be a *native-norm* for ordering these acoustic signals as evidenced by the significant correlations. Although speakers with different L1 sound systems encounter different problems when trying to acquire native-like stress production, encouragingly, it appears that they may be able to learn through increased L2 input. Not only do experienced learners produce fewer pronunciation errors, they also produce ELS in a more native-like manner. For example, although Saudi speakers inherently under-use duration as an acoustic cue to ELS by not fully reducing unstressed vowels, they are able to use this acoustic correlate more accurately as their language skills progress. Similarly, Chinese learners of English are able to overcome the negative transfer of their tonal system by employing pitch in a more native-like manner as they advance in their studies.

While this project has already yielded interesting results, it is a work in progress for my Master's thesis, and there are still many more research questions that can potentially be answered using the current dataset. These include issues relating to conflicts with standard dictionary pronunciations, the role of tonic accent shift, the idiosyncrasies of individual suffixes, and L1-specific acoustic correlates of stress salience hierarchies.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the assistance of Professors Ettien Koffi and Choon Kim for their invaluable advice and assistance. Also, many thanks to the Statistical Consulting and

Research Center at St. Cloud State University (SCSU). This research was partially funded by the Student Research Fund awarded by SCSU's Office of Research and Sponsored Programs.

ABOUT THE AUTHOR

Paul Keyworth has taught EFL/ESL since 2002 in a wide variety of educational contexts in South Korea, Singapore, the UK, and the USA. Currently, he is pursuing his Masters in TESL with an emphasis in applied linguistics. His special interests include pedagogical grammar, academic writing, acoustic phonetics, and computer-assisted pronunciation training. He is also interested to learn more about the emerging field of Forensic Acoustics.

REFERENCES

Altmann, H. (2006). The perception and production of second language stress: A cross-linguistic experimental study. *Dissertation Abstracts International, Section A: The Humanities And Social Sciences*, *67*(6), 2133-2134. (2006753642)

Bouchhioua, N. (2008). Duration as a cue to stress and accent in Tunisian Arabic, native English, and L2 English. *Proceedings from Speech Prosody 2008: The Fourth International Conference on Speech Prosody*. (pp. 535-538). Campinas: Brazil.

Burzio, L. (1994). *Principles of English stress: Cambridge Studies in Linguistics #72, 1st edition*. Cambridge: Cambridge University Press.

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

de Jong, K., & Zawaydeh, B. A. (1999). Stress, duration, and intonation in Arabic word-level prosody. *Journal of Phonetics, 27,* 3-22.

Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing (1st edition)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Flege, J. E., & Bohn, O. S. (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition, 11*, 35-62.

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

Fry, D. B. (1958). Experiments in the perception of stress. *Language & Speech*, *1:2,* 126-152.

Kreidler, C. W. (2004). *The pronunciation of English: A course book, 2nd edition.* Malden, MA: Blackwell.

Ladefoged, P. (2001). *A course in phonetics*, 4th edition. New York: Thomson-Wadsworth.

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

Ladefoged, P., & Johnson, K. (2010). *A course in phonetics*, 6th edition. New York: Thomson-Wadsworth.

Lee, S., & Cho, M. (2011). An acoustic analysis of stressed and unstressed vowels in English nouns: A cross-language case study. *Korean Journal of English Language and Linguistics 17,* 61-88.

Maeda, S. (1976): A characterization of American English intonation. (Unpublished doctoral dissertation). MIT: Cambridge, MA.

McClean, M.D., & Tiffany, W.R. (1973). The acoustic parameters of stress in relation to syllable position, speech loudness and rate. *Language and Speech, 16*(3), 283-290.

Munro, M. & Derwing, T. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*, 265-292.

Ryan, K. (2005). Grid-maker.praat [Praat script]. Available from http://www.linguistics.ucla.edu/faciliti/facilities/acoustic/praat.html

Shemshadsara, Z.G. (2011). EFL learner's awareness of stress-moving vs. neutral suffixes. English Language Teaching, 4:4, 146-153.

Yoon, T. (2008). Stress analysis script [Praat script]. Available from http://web.uvic.ca/~tyoon/resource/vq.praat

Zhang, Y., Nissen. S.L., & Francis, A.L. (2008). Acoustic characteristics of English lexical stress produced by Mandarin speakers. *Journal of the Acoustical Society of America, 123:6,* 4498-4513.

Zuraiq, W. (2006). The production of lexical stress by native speakers of Arabic and English and by Arab learners of English. *Dissertation Abstracts International, Section A: The Humanities And Social Sciences*, *66*(12), 4375. (2006751445)

Appendix A

Table A1

Ladefogedian stress-pattern values of the stimuli

Token/ Variable	CV					Stress-pattern Value
1.<history>	[hɪ́s]	[tri]				**1 2**
Stress	+	+				
Tonic accent	+	-				
Full vowel	+	+				
2.<historic>	[hɪs]	[tɔ́r]	[ɪk]			**2 1 3**
Stress	+	+	-			
Tonic accent	-	+	-			
Full vowel	+	+	+			
3.<historical>	[hɪs]	[tɔ́r]	[ɪ]	[kəl]		**2 1 3 4**
Stress	+	+	-	-		
Tonic accent	-	+	-	-		
Full vowel	+	+	+	-		
4.<historicity>	[hɪs]	[tə]	[rɪ́s]	[ɪ]	[ti]	**2 4 1 3 3**
Stress	+	-	+	-	-	
Tonic accent	-	-	+	-	-	
Full vowel	+	-	+	+	+	
5.<historial>	[hɪs]	[tɔ́r]	[i]	[əl]		**2 1 3 4**
Stress	+	+	-	-		
Tonic accent	-	+	-	-		
Full vowel	+	+	+	-		
6.<historify>	[hɪs]	[tɔ́r]	[ɪ]	[faɪ]		**2 1 3 3**
Stress	+	+	-	-		
Tonic accent	-	+	-	-		
Full vowel	+	+	+	+		
7.<historious>	[hɪs]	[tɔ́r]	[i]	[əs]		**2 1 3 4**
Stress	+	+	-	-		
Tonic accent	-	+	-	-		
Full vowel	+	+	+	-		
8.<historian>	[hɪs]	[tɔ́r]	[i]	[ən]		**2 1 3 4**
Stress	+	+	-	-		
Tonic accent	-	+	-	-		
Full vowel	+	+	+	-		

Note. Adapted from "Teaching Pronunciation: A Course in Phonetics," by P. Ladefoged, 2001, p97.

- Phonetic transcriptions based on International Phonetic Alphabet (IPA) pronunciations provided by Dictionary.com (2013).

Appendix B

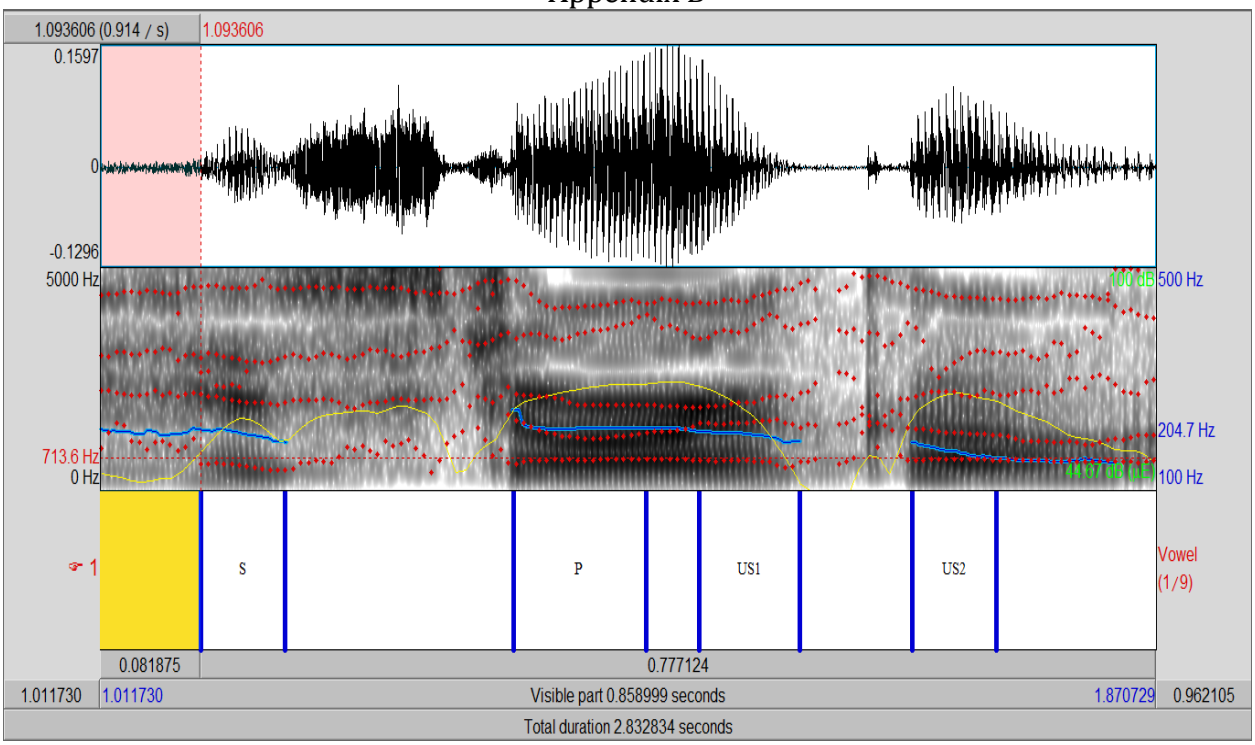


Figure B1. <Historical> as spoken by a Midwest American female.
P = primary-stressed vowel *S* = secondary-stressed vowel *US* = unstressed vowel

Makino, T. (2014). Pronunciation characteristics of Japanese speakers' English: A preliminary corpus-based study. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 121-136). Ames, IA: Iowa State University.

PRONUNCIATION CHARACTERISTICS OF JAPANESE SPEAKERS' ENGLISH: A PRELIMINARY CORPUS-BASED STUDY

Takehiko Makino, Chuo University, Tokyo

The development of the English Read by Japanese (ERJ) Phonetic Corpus consists of computer-readable narrow phonetic transcriptions and their corresponding target phonemes of selected 800 utterances from ERJ speech database. In describing the pronunciation characteristics of English spoken by Japanese speakers (or speaker of any language), we have been relying on the "rules of thumb" based on informal observations or theoretical predictions from the L1-L2 phonological differences, such as L/R confusion or conflation of English vowels into a five-vowel system in the case of Japanese speakers. While such rules of thumb have had roles to play, corpus-based studies of other areas of linguistic research have proved that they cannot give us the total picture of what are being studied, and L2 pronunciation should not be an exception. Indeed, a preliminary survey of the ERJ Phonetic Corpus has revealed some rather unexpected findings. The most notable of such findings is the spirantization (fricative realization) of voiceless plosives. Such a process is not part of standard Japanese phonology and cannot be the case of a negative L2 transfer. We can expect that the Corpus will help make a more systematic description of Japanese speakers' pronunciation of English.

Background

The purpose of this paper is to make an interim report on the findings from a corpus-based descriptive study of the characteristics of Japanese speakers' pronunciation of English. I am currently developing English Read by Japanese (ERJ) Phonetic Corpus (Makino, 2007, 2009; Makino & Aoki 2012). The Corpus consists of computer-readable narrow phonetic transcriptions and their corresponding target phonemes and words of selected 800 utterances from ERJ speech database, a large collection (more than 70,000 speech files) of read-aloud sentences by Japanese university students.

The rationale for the corpus development was the lack of a systematic survey of the characteristics of Japanese speakers' pronunciation of English. Corpus studies on L2

pronunciation have been very rare (Gut, 2009; Li, Zhang, Li, Harrison, Lo, & Meng, 2011; Meng, Tseng, Kondo, Harrison, & Viscelgia, 2009). The Corpus is intended to help fill this gap.

Introduction to ERJ speech database

ERJ stands for "English Read by Japanese" and the database was collected mainly in order to help CALL system development (Minematsu, et al. 2002). 807 different sentences and 1,009 different words (or word sets) were read aloud by 200 (100 male and 100 female) speakers in 20 different recording sites in Japan. All of the sites were universities and all the speakers were students there. Probably because it was deemed unpractical to ask all the 200 speakers to record all the sentences and words, they were divided into several sets, and individual speakers were asked to record only one sentence set and one word set. As a result, each sentence was read aloud by about 24 speakers (12 males and 12 females) and each word by about 40 speakers (20 males and 20 females). In total, the ERJ speech database consists of more than 70,000 speech files: 24,744 sentence files and 45,495 word files.

Training sheets

Before recording, the speakers were asked to practice pronouncing the words and sentences with training sheets which presented phonemic notations as well as orthographic words and sentences. The phonemic symbols used in the training sheets are based on ARPAbet used in TIMIT database (Garofolo, et al. 1993) and the CMU Pronouncing Dictionary (Carnegie Mellon University 2008), listed below:

Consonants: P, T, K, B, D, G, CH, JH, F, TH, S, SH, HH, V, DH, Z, ZH, M, N, NG, W, Y, L, R

Vowels: IY, IH, EY, EH, AE, AA, AO, OW, UH, UW, AH, AX, AW, AY, OY, ER, AXR

The model of the pronunciation is therefore mainstream American English. Each vowel was specified for degrees of stress: "1" for primary, "2" for secondary and "0" for unstressed.

In order to ensure that the speakers understood these symbols correctly, a website was prepared where they could listen to word examples of each phonemic symbol.

Examples from training sheets:

S1_001 This was easy for us.
[DH IH1 S] [W AA1 Z] [IY1 Z IY0] [F AO1 R] [AH1 S]
S1_002 Is this seesaw safe?
[IH1 Z] [DH IH1 S] [S IY1 S AO2] [S EY1 F]
S1_003 Those thieves stole thirty jewels.
[DH OW1 Z] [TH IY1 V Z] [S T OW1 L] [TH ER1 T IY0] [JH UW1 AX0 L Z]

The recordings

In the recording sessions, the scripts only presented orthographic words and sentences, without phonemic notations. The reason for the change was that reading phonemic notations can induce unnatural pronunciation.

Corpus building procedure
Limitations of the ERJ data

It follows from the nature of the recording procedure above that there are some limitations in the ERJ speech database:

a) The speech is not spontaneous but read-aloud. It does not represent what is happening in natural settings.
b) The sentences are all isolated and out-of-context. This could have led to improper prosody (accenting, intonation or rhythm).
c) The words in the TIMIT phonologically-balanced sentences were chosen for their sounds. As a result, many of them were unfamiliar to the subjects. Even though phonemic notations were presented at the training stage, this could have led to mispronunciations or awkward pronunciations.

Choice of the materials used

Obviously, it is absolutely unpractical to use the whole database for the corpus building because of its sheer size. I have chosen to transcribe the 800 sentence files used in another study (Minematsu, Okabe, Ogaki, & Hirose, 2011). In that study, the recordings were played over the telephone to Americans who were not familiar with Japanese speakers' English, who then repeated what they (thought they) heard, and those repetitions were transcribed orthographically.

These sentences are all from the phonologically-balanced sets of sentences. The exclusion of word sets is justified because we are not interested in the pronunciation of individual words.

Transcription procedure

The transcription procedure for ERJ Phonetic Corpus is listed below:

1) To reduce the effort of manual transcription, the files were pre-processed by the Penn Phonetics Lab Forced Aligner (Yuan & Liberman, 2008; http://www.ling.upenn.edu/phonetics/p2fa/), which produced forced aligned transcriptions of English words and phonemes for each file in the Praat (Boersma & Weenik, 2013) TextGrid format.
2) Then, using Praat, the TextGrids were re-formatted into three tiers (target words, target phones, actual phones). The actual phones were manually transcribed, and boundaries of target phones and target words were manually aligned with those of the actual phones.
3) The corrected TextGrids were then imported into ELAN (Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006; http://tla.mpi.nl/tools/tla-tools/elan/), which has much better searching functionality than Praat, allowing searching in multiple files and output in concordance format.

The resulting ELAN .eaf files and the original .wav files are the complete individual data of the Corpus.

The data set for this study

The ERJ Phonetic Corpus consists of 419 different sentences, read by 200 different speakers (100 males and 100 females). Most of the sentences were read by two (one

male and one female) speakers. The total number of words is 5,959, with 1,599 different words. The total number of target phonemes is 24,873. The number of different target phonemes is, of course, 41 (the number of phonemes in General American English. cf. §2.2). The total number of actual phones is 25,460. The total number of different actual phones is 481.

Table 1

The tokens of each phoneme in ERJ Phonetic Corpus

Vowels				Consonants		
ARPAbet	IPA	Count		ARPAbet	IPA	Count
IY	i	900		P	p	630
IH	ɪ	943		B	b	540
EY	eɪ	448		T	t	1510
EH	ɛ	605		D	d	959
AE	æ	640		K	k	901
AA	ɑ	605		G	g	339
AO	ɔ	453		F	f	480
OW	oʊ	345		V	v	387
UH	ʊ	131		TH	θ	144
UW	u	512		DH	ð	565
AH	ʌ	372		S	s	1223
AX	ə	2283		Z	z	868
AY	aɪ	416		HH	h	266
OY	ɔɪ	85		SH	ʃ	247
AW	aʊ	153		ZH	ʒ	27
ER	ɝ	200		CH	tʃ	209
AXR	ɚ	546		JH	dʒ	252
				M	m	718
				N	n	1496
				NG	ŋ	258
				L	l	1203
				R	r	1257
				W	w	453
				Y	j	294

Predictions about vowels

Japanese has a five-vowel system: /i, e, a, o, u/, whose typical realizations are [i, e, ɐ, o, ɯ]. Based on this, the following predictions can be made about the realizations of English vowel phonemes by Japanese speakers:

IH, IY	[i(ː)]
UH, UW	[ɯ(ː)]
AA, AE, AH, ER, AXR, AX	[ɐ(ː)]
EH, EY	[e(ː)]
AO, OW	[oː]
AY, OY, AW	combinations of /a, i, o/: [ɐi], [oi], [ɐɯ]

Findings about vowels - IY /i/ and IH /ɪ/

Before looking at the findings, I have to note that the data below represent only the one-to-one relationship between target phonemes and actual phones, due to the limitation in the searching function of ELAN. As a result, they cannot show a complete picture, which will eventually have to be addressed.

The realizations of English IY /i/ and IH /ɪ/ are shown in the following table, arranged by frequency:

Table 2

The frequency of different phones for target /i, ɪ/ in ERJ Phonetic Corpus

IY /i/	**Count (N=874)**		**IH /ɪ/**	**Count (N=905)**
i	613		i	445
ï	98		ɪ	287
ɪ	66		ï	48
ĩ	19		ĩ	30
i̥	16		i̥	24
e	15		ɐ	13
others	47		ə	10
			others	48

More than half of the realizations of IH are [i]-like, which is also the dominant realization of IY. This is probably due to negative L1 transfer. On the other hand, we

found a substantial number of IH's realized as [ɪ]. So the conflation of these phonemes is not complete.

The [ɪ] in IY is interesting because the Japanese language does not have this phone as an allophone of any of its phonemes. It may be a case of hypercorrection.

UW /u/ and UH /ʊ/

The realizations of English UW /u/ and UH /ʊ/ are shown in the following table, arranged by frequency:

Table 3

The frequency of different phones for target /u, ʊ/ in ERJ Phonetic Corpus

UW /u/	**Count (N=483)**		**UH /ʊ/**	**Count (N=90)**
ʉ	163		ʉ	23
ʉ̈	102		ʉ̈	19
ɨ	86		ʊ	18
ü	75		ü	8
ɤ	9		ɨ	7
ʊ	9		other	15
other	39			

Although the tokens of UH may not be sufficient, we do not see any difference in the distributions of different realizations of UW and UH. This is probably a case of negative L1 transfer.

AA /ɑ/, AE /æ/, AH /ʌ/, ER /ɝ/, AXR /ɚ/ and AX /ə/

The realizations of English AA /ɑ/, AE /æ/, AH /ʌ/, ER /ɝ/, and AX /ə/ are shown in the following table, arranged by frequency:

Table 4

The frequency of different phones for target /a, æ, ʌ, ɝ, ɚ/ in ERJ Phonetic Corpus

AA	**Count**	**AE**	**Count**	**AH**	**Count**	**ER**	**Count**	**AXR**	**Count**

/ɑ/	(N=436)	/æ/	(N=623)	/ʌ/	(N=362)	/ɝ/	(N=194)	/ɚ/	(N=517)
ɐ	94	ɐ	329	ɐ	169	ɐ	105	ə	172
ö	58	a	68	ə	42	ə	29	ɐ	158
ɔ	53	ə	63	ɐ̃	30	ɚ	27	ɚ	84
ɔ̈	44	æ	60	ö	25	etc	33	ɨ	14
o	39	ɐ̃	37	a	19			etc	89
ə	32	etc	66	ɵ	11				
ɵ	32			etc	66				
etc	84								

It is evident that most of the phonemes here are realized as a phone typically representing the Japanese phoneme /a/.

The o-like realizations for AA have occurred because of the influence from spelling. AA is typically represented with <o>, which is used in Romanized representation of the Japanese /o/, e.g. *tori* ‘bird’ /tori/ [toɾi].

The realizations of English AX (schwa) are shown in Table 5, arranged by frequency. Although the phones [ə, ɐ] rank the highest, the realizations of schwa is much more diverse than other a-like phonemes. This again may reflect its spelling, because schwa is represented with various spellings in English.

Table 5

The frequency of different phones for target /ə/ in ERJ Phonetic Corpus

AX /ə/	**Count (N=2081)**									
ə	495		ĩ	58		ö	18		others (ɔ̃, eɪ, ɪ̰, ʏ, ɑ̤, ɛ̃, ɜ, ʌ, y, ə̰, əɚ, əɪ, ɛɪ, ḭ, ʔɐ, ä, ɒ, ɐ̰, ɐ̰, ɑˤ, ɐðɨ, æ̃, ɐi, ɐɪ, ɐ̃ɪ, ẽ, əd̄, əɪ̰, əɯ, ən, ət, hɐ̰, ɦɐ, i̤, ɪ̰, ɪ̃, ɪʉ, l, ɯ, ɯ̃, ɯ̈, ø, ötʊ, ɸ, ɹɪ, s, ü, ũ, ʉ̃, ʊ, ÿ, z̃, ʔə)	94
			ɵ	46		o	16			
ɐ	347		ẽ	34		ɪ̃	14			
			ɛ	33		ɯ̈	14			
i	225		a	27		ɨ̃	12			
			ə̃	23		ɔ̈	11			
ɪ	192		ʉ	22		ɛ̈	11			
			õ	20		æ	10			
ɨ	136		ɵ̃	20		ɔ	10			
			ɐ̃	19		ɤ	10			
e	108		ë	19		ɚ	9			
			ï	19		ḭ	9			

Predictions about consonants

Japanese has the following consonant phonemes: /p, t, k, b, d, g, s, z, h, m, n, r, j/. It also has special moraic phonemes which are usually realized as consonants: a moraic nasal /N/ and a moraic obstruent /Q/. Some of the realizational rules of the consonantal phonemes are:

a) Plosives /p, t, k, b, d, g/: voiced set tends to be realized as fricatives (spirantized) between vowels; /g/ can also be realized as [ŋ] in the same environment.
b) Fricatives /h, s, z/: /h/ is realized as [ç] before /i/ and [ɸ] before /u/; /z/ is often realized as an affricate [dz] word-initially.
c) Dental/alveolar sounds /t, d, s, z, n/ are heavily palatalized before /i/ and /j/ and realized as [tɕ, (d)ʑ, ɕ, (d)ʑ, ɲ] respectively, although non-palatalized pronunciation is possible for loanwords.
d) A liquid /r/: typically a tap [ɾ], but [l, ɖ] are also possible.
e) A moraic nasal /N/ appearing syllable-finally; a nasal stop homorganic to the following consonant; a nasalized vowel before a vowel; when absolute final a uvular nasal [ɴ].
f) No phonemic consonant cluster; no syllable final consonants, except for the moraic nasal /N/ and obstruent /Q/.

Consonants: Some possible predictions

Based on the above, the following predictions, among others, can be made about the realization of English consonant phonemes:

a) Lack of distinction between /l/ and /r/ before vowels
b) Lack of distinctions between:
   /s/ and /θ/
   /z/ and /ð/ and /dz/
   /dʒ/ and /ʒ/
   /b/ and /v/
   /h/ and /f/ before [u]-like phonemes
c) Addition of a weak vowel ([ɯ, ɨ, ə]) after a word-final consonant or in consonant clusters

Unexpected findings: Spirantization of a voiceless plosive /p/

The realizations of English /p/ are shown in the table below. The list is categorized according to the phonological contexts and arranged by frequency:

Table 6

The frequency of different phones for target /p/ in different phonological contexts in ERJ Phonetic Corpus

	Phonological contexts (sil < silence)								
realization	**V_C**	**C_V**	**V_V**	**C_C**	**sil_C**	**sil_V**	**V_sil**	**C_sil**	**Total**
p	25	94	74	39	12	15			259
ɸ	49	12	18	22	8	5	6	2	122
pʰ	19	12	21	10	7	4	10	1	84
pɸ	8	1	1	5	1		3	1	20
pɨ	7		1	5	2				15
p̚	12						2		14
other	10	2	4	13	3	1	1	0	34
Total	**130**	**121**	**119**	**94**	**33**	**25**	**22**	**4**	**548**

Vowel insertion before another consonant is what theories predict, but it is hardly the most frequent case. We find numerous cases of spirantized phones. They cannot be predicted from Japanese phonology, where voiceless plosives are not spirantized.

Spirantization is most likely to occur syllable-finally, especially before consonants. But note that it occurs prevocalically as well, which is not the most likely position for consonantal weakening. This fact can have repercussions to phonological theories.

Unexpected findings: Spirantization of a voiceless plosive /k/

The realizations of English /k/ are shown in the table below. The list is categorized according to the phonological contexts and arranged by frequency:

Table 7

The frequency of different phones for target /k/ in different phonological contexts in ERJ Phonetic Corpus

	Phonological contexts (sil < silence)								
realization	**V_C**	**V_V**	**C_V**	**C_C**	**sil_V**	**V_sil**	**sil_C**	**C_sil**	**Total**
k	80	100	87	23	30	1	8		329
kʰ	76	25	35	26	17	20	4	1	204
k̟	1	17	20	4	11				53
x	36	4		1		2	1		44
kɨ	9	4		6			4		23
k̟ʰ	1	11	5		4				21
k͡	16								16
k’	6	3		1		4			14
kx	6			2		3		1	12
others	23	7		16		2	5		53
Total	**254**	**171**	**147**	**79**	**62**	**32**	**22**	**2**	**769**

Here the spirantization is less frequent than in the case of /p/, but the most likely contexts are the same. Spirantization is not at all frequent in /t/, which is not addressed here.

L and R

The realizations of English /l/ and /r/ are shown in the table below. The list is categorized according to the phonological contexts and arranged by frequency:

Table 8

The frequency of different phones for target /l, r/ in different phonological contexts in ERJ Phonetic Corpus

	Phonological contexts (sil < silence)								**Phonological contexts (sil < silence)**					
/l/	**C_V**	**V_C**	**V_V**	**V_sil**	**sil_V**	**Total**		**/r/**	**C_V**	**V_V**	**V_C**	**sil_V**	**V_sil**	**Total**
l	134	102	119	20	9	384		ɹ	303	61	3	22		389
ɬ	32	109	32	30		203		ɾ	120	30		12		162
ɾ	79	2	50		8	139		ə	1	4	54		10	69
ɹ	58		41	2	1	102		l	33	17		7		57
lɨ		25	4	8		37		ɚ	1	7	26		6	40
ɾɨ		27	1	8		36		ɐ		3	20		13	36
others	43	47	23	20	2	135		ɯ̟	28	1				29
								ɻ	20	5		2		27
								others	32	5	20	7	5	69
Total	**346**	**312**	**270**	**88**	**20**	**1036**		**Total**	**538**	**133**	**123**	**50**	**34**	**878**

Japanese speakers' English /r/ is notoriously mispronounced, but the data shows that it is not that bad. Nearly half the occurrence of this phoneme conforms to the native-speaker target. [ɾ] is what theories predict in both phonemes, but it is not actually in the majority.

/l/ is probably the more problematic phoneme for Japanese speakers of English. Note the case of [ɹ] for /l/. This could be the case of hypercorrection, probably because /r/ is one of the most emphasized sound in the teaching of the pronunciation of English in Japan, whereas /l/ is hardly ever emphasized.

Implications for pronunciation teaching

The study presented here is preliminary and incomplete, but I believe that I have shown that the reality is far more complicated than what phonological theory predicts.
In Japan, instruction of pronunciation of English at an introductory level is patchy at best. /r/ is emphasized but /l/ is not, /θ/ and /ð/ but not /s/ and /z/, vowels are not taught as an overall system but only some of the individual vowels deemed difficult such as /æ/

and /ɚ/ are treated, if at all. This may have resulted in the sort of disparate pronunciations presented above. We have to study what is happening more closely and help make a better pronunciation teaching syllabus.

Prosodic notation

So far only the segmental transcription has been completed for ERJ Phonetic Corpus. Since the Corpus is intended to be a source of all the phonetic characteristics of Japanese speakers' English speech, prosodic transcription is also necessary.

L2 prosody is very difficult to describe. Studies of non-native prosody such as Gut (2009) and Li, et al. (2011) use (modified) English ToBI, which I think is a wrong thing to do. L2 prosodic system is neither that of L1 nor of the target language, but something of the mixture of the two. I am now addressing this problem and devising a notational system of Japanese speakers' English prosody.

Mispronunciation and misperception

As noted above, the data set in this study was the same as that used in Minematsu, et al. (2011), where the recordings were played over the telephone to Americans who were not familiar with Japanese speakers' English and they were asked to repeat the sentences they heard. Their responses were orthographically transcribed.

With this data, we will be able to explore what sort of actual phone deviations are likely to lead to misunderstandings of what sort. This can be the basis for a study of intelligibility.

ACKNOWLEDGEMENTS

The research for this study was supported by Grant-in-Aid for Scientific Research (B) No.23300067 (project leader: Nobuaki Minematsu) from the Japan Society for the Promotion of Science, and a Chuo University Overseas Research Program. The latter enabled me to conduct most of the work for this paper as a Visiting Scholar at the University of Pennsylvania and the Linguistic Data Consortium.

ABOUT THE AUTHOR

Takehiko Makino is an Associate Professor of English as a Foreign Language at Chuo University (Tokyo, Japan) and has taught English phonetics at a number of other universities including his alma mater Tokyo University of Foreign Studies, from which he received an MA in English linguistics in 1991. He also studied linguistics at the University of Kansas, USA in 1987-88. In 2004, he was awarded a Certificate of Proficiency in the Phonetics of English by the International Phonetic Association. From April 2012 till March 2014, he had an appointment as a Visiting Scholar at the University of Pennsylvania and the Linguistic Data Consortium under Chuo University Overseas Research Program.

Address:
Chuo University
742-1 Higashi Nakano
Hachioji-shi, Tokyo 192-0393, Japan
Telephone: +81-(0)42-674-3401
Email: mackinaw@tamacc.chuo-u.ac.jp

REFERENCES

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer (version 5.3.55) [Computer software]. Available: http://www.praat.org/

Carnegie Mellon University (2008). CMU Pronouncing Dictionary (v. 0.7a). [Electronic database]. Available: http://www.speech.cs.cmu.edu/cgi-bin/cmudict

Garofolo, John, Lori Lamel, William Fisher, Jonathan Fiscus, David Pallett, Nancy Dahlgren, & Victor Zue. (1993). TIMIT Acoustic-Phonetic Continuous Speech Corpus LDC93S1. Web Download. Philadelphia: Linguistic Data Consortium. [Electronic database]. Available: https://catalog.ldc.upenn.edu/LDC93S1

Gut, U. (2009). *Non-native speech: A corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfurt: Peter Lang

Li, M., Zhang, S., Li, K., Harrison, A., Lo, W.-K., & Meng, H. (2011). Design and collection of an L2 English corpus with a suprasegmental focus for Chinese learners of English. *Proceedings from the 17th International Congress of Phonetic Sciences*, Hong Kong, 1210-1213.

Makino, T. (2007). A corpus of Japanese speakers' pronunciation of American English: preliminary research. Paper presented at Phonetics Teaching and Learning Conference 2007. Abstract retrieved from http://www.ucl.ac.uk/psychlangsci/ptlc/proceedings_2007/ptlcpaper_02e

Makino, T. (2009). Vowel substitution patterns in Japanese speakers' English. In Biljana Čubrović & Tatjana Paunović (Eds.), *Ta(l)king English Phonetics Across Frontiers* (pp.19-31). Newcastle: Cambridge Scholars Publishing.

Makino, T., & Aoki, R. (2012). English Read by Japanese Phonetic Corpus: An interim report [Electronic version]. *Research in Language*, 10(1), 79–95.

Meng, H., Tseng, C., Kondo, M., Harrison A., & Viscelgia, T. (2009). Studying L2 suprasegmental features in Asian Englishes: a position paper. *Proceedings from Interspeech 2009*, Brighton, UK, 1683-1686.

Minematsu, N., Okabe, K., Ogaki, K. & Hirose, K. (2011). Measurement of objective intelligibility of Japanese accented English using ERJ (English Read by Japanese) database. *Proceedings from Interspeech 2011*, Florence, Italy, 1481-84.

Minematsu, N., Tomiyama, Y., Yoshimoto, K., Shimizu, K., Nakagawa, S., Dantsuji, M., & Makino, S. (2002). English speech database read by Japanese learners for CALL system development. *Proceedings from the 3rd International Conference of Language Resources and Evaluation*, Las Palmas, Spain, 896-903.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality Research. *Proceedings from the 5th International Conference of Language Resources and Evaluation*, Genoa, Italy, 1556-59.

Yuan, J., & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *Proceedings from Acoustics '08*, Paris, France, 5687-5690.

McCrocklin, S. , & Link, S. (2014). What is identity? ELL and bilinguals' views on the role of accent. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 137-144). Ames, IA: Iowa State University.

WHAT IS IDENTITY? ELL AND BILINGUALS' VIEWS ON THE ROLE OF ACCENT

Shannon McCrocklin, University of Texas Pan American
Stephanie Link, Iowa State University

Many researchers and theorists have proposed a connection between accent and identity (Celce-Murcia, Brinton, Goodwin, & Griner, 2010; Ochs, 1993; Setter and Jenkins, 2005). Some, however, have gone beyond this, indicating that students fear obtaining a native speaker accent. "To speak an L2 like a native is to take a drastic step into the unknown, accompanied by the unconscious fear of no return..." (Daniels, 1995, p. 6). Yet, this comment may strike many teachers and researchers as surprising because as Sobkowiak points out, "in my whole teaching career I have not met a [student] who would not like to sound like a native, or who would fear to step on this 'road of no return'" (2005, p. 144). Perhaps the difference in perspective can be explained by a difference in language learning group, those that have successfully spoken English like a native and those that have not. This research study examines the perceptions of English Language Learners (ELLs) and English speaking bilinguals. Subjects participated in a semi-structured interview to discuss their experiences interacting with others and their perceptions of accent and identity. Findings from the interviews suggest that these two groups have very different ideas about accent and identity. Results from this study provide insights into the possible misconceptions and assumptions that underscore our work as educators and researchers and can hopefully be used to inform future teaching in the field of pronunciation.

INTRODUCTION

Many researchers and theorists have proposed a connection between accent and identity, not only for first language accent but also for second language or foreign accents. Some, however, have gone beyond this, indicating that students fear obtaining a native speaker accent because they might lose part of their identity. Yet, previous research seems to indicate the students overwhelmingly want to obtain a native accent. Research is needed to see the issues of identity, accent, and fear not only from the perspective of English Language Learners (ELLs), but also from the perspective of bilinguals who have successfully obtained a native accent in English.

Background Information

Identity, "a range of social personae, including social statuses, roles, positions, relationships, and institutional and other relevant community identities one may attempt to claim or assign in the course of social life" (Ochs, 1993, p. 288), is a complex issue due to its non-static, negotiated nature (Marx, 2002). An important component of this definition is the construction of identity through social interaction (see Haslam, 2001 for a discussion of the social identity approach), suggesting that identity is shaped through overlapping circles of internalized group memberships. That is, social identity is not one personal self, but rather by individual-based perception and

self-knowledge of how one relates to various groups, enabling the development of multiple social identities.

Many researchers and theorists have proposed a connection between accent and identity (Celce-Murcia, Brinton, Goodwin, & Griner, 2010; Ochs, 1993; Setter and Jenkins, 2005). This link has been discussed not only for first language accent (Levon, 2006; Johnstone & Kiesling, 2008), but also for second language or foreign accents (Marx, 2002; Rindal, 2010). According to Derwing and Munro (1997), a foreign accent indicates that the sounds are being produced in a way outside of the norm of English. This is not the same as intelligibility, which refers to the ability to understand the words and phrases being spoken. Despite prevailing beliefs that a strong accent would lead to intelligibility issues, Derwing and Munro point out that research has shown that a person can have a strong and noticeable accent without losing intelligibility. This is supported by other research as well (Munro & Derwing, 1995; Flege, Takagi, & Mann, 1995; Munro, Flege, & MacKay, 1996). For example. a native accent could represent identification with or memberships in a group (Celce-Murcia, Brinton, & Goodwin, 2010) or a foreign accent could be a signal that someone is different (Bresnahan, Ohash, Nebashi, Liu, & Shearman, 2002).

Some, however, have gone beyond this, indicating that students fear obtaining a native speaker accent. "To speak an L2 like a native is to take a drastic step into the unknown, accompanied by the unconscious fear of no return..." (Daniels, 1995). In terms of teacher responsibility, Porter and Garvin (1989) further argue:

> …A person's pronunciation is one expression of that person's self-image. To seek to change someone's pronunciation—whether of the L1 or of an L2—is to tamper with their self-image, and is thus unethical—morally wrong. (p. 8, as cited in Dalton & Seidlhofer, 1994)

Yet, these attitudes may strike many teachers and researchers as surprising because as Sobkowiak points out, "in my whole teaching career I have not met a [student] who would not like to sound like a native, or who would fear to step on this 'road of no return'" (2005). Research supports the notion that most students want to sound like native speakers (Andreasson, 1994; Derwing, 2003; Yamaguchi, 2002). For example, Derwing (2003) found that 95% of ELL immigrant respondents in Canada would choose to speak like native speakers if they could.

Previous research into language learning goals, however, has primarily looked at the goals of current language learners. Perhaps the difference in perspective can be explained by a difference in language learning group, those who have successfully spoken English like a native (bilinguals) and those who have not. For the purpose of this study, it is important to note that we are defining a bilingual as a person who, due to having learned two languages from childhood, is fully competent in both languages. Further, regarding pronunciation, "fully competent" is considered to mean having a native or native-like accent. Bilinguals may present a unique view of accent and identity as they often have two very different cultural groups that they interact with. Some of the previous research with bilinguals has examined their expression of dual identities through code-switching (Auer, 2002) or how allegiance to one particular group can affect the view of the usefulness of being bilingual (Norton, 1997). To our knowledge, no studies have been conducted to examine bilinguals' attitudes towards native accents or fear of identity loss due to a native accent.

Research Questions

To examine the links between accent and identity further, we wanted to examine the question, "Do bilinguals view the relationship between accent and identity the same as ELL students?" Specifically, we wanted to address the two following research questions:

1. What attitudes and expectations do ELLs and bilinguals hold toward native-like accents?
2. How do ELLs and bilinguals perceive a link between their accent in the L2 to their identity?

Findings from this study hope to provide insights into the possible misconceptions and assumptions that underscore our work as educators and researchers and can hopefully be used to inform future teaching in the field of pronunciation.

METHODS

To examine the attitudes of both the ELLs and bilinguals, a basic interpretive qualitative research design (Merriam, 2002) was employed. In depth semi-structured interviews were conducted to examine perceptions of the connection between accent and identity and its influence on language learning goals.

Participants

The study took place at a major culturally diverse, mid-western university in the United States located in a primarily monolingual, white, medium-sized city. All participants were undergraduate students at the university. Eight ELL students participated, seven Chinese students and one Kuwaiti. All of the ELL students had lived in the U.S. for less than two years and had a noticeable foreign accent in English. Five bilinguals participated, two Mexicans, one Albanian, one Nigerian, and one Vietnamese. All of the bilinguals had lived in the U.S. since early childhood and lacked a noticeable foreign accent in English.

Interviews

To examine participants' beliefs about the link between accent and identity, participants were interviewed for 30-40 minutes. The interview was semi-structured with 21 basic questions. The first 16 questions were about language background (adopted from Shin, 2010). The final five questions asked about accent and identity. Some of these questions had to be slightly altered to fit the language learning backgrounds of each group, but the questions were paired in meaning. For example, the ELLs were asked, "If you could sound like a native speaker, like right now wave my magic wand, would you take that native accent?" whereas the bilinguals were asked, "If you could add an accent to your English speech right now, one that would show where you are from, would you take it?" Where needed, interviewers paraphrased questions if the participants seemed to struggle to understand the initial question asked. The interviewers also followed-up on questions as needed to get more information and details.

Responses were audio recorded and transcribed verbatim. Each individual researcher first did an exploratory analysis, coding for patterns and themes related to the research questions, as well as

a confirmatory analysis, coding for indications of a link between accent and identity. Finally, the researchers peer debriefed, checking emerging themes and conclusions.

RESULTS - ATTITUDES AND EXPECTATIONS

To understand students' attitudes towards a native accent and answer the research question, "What attitudes and expectations do ELLs and bilinguals hold toward native-like accents?" the ELLs were asked, "If you could sound like a native speaker, like right now wave my magic wand, would you take that native accent?" All of the students said yes. Jing, a Chinese L1 student, stated in response the following:

Jing: Yeah.
Interviewer: Ok Why?
Jing: Why not?

To distinguish between issues of accent and intelligibility, students were then asked, "If starting today you would never be misunderstood again, would you still want to sound like a native speaker?" All students retained their previous position, saying yes. One Chinese student said, "Actually, if I have a native speaking accent, I'm very proud…"

The bilinguals were asked the converse question "If you could add an accent to your English speech right now, one that would show where you are from, would you take it?" All students said no. One student said, "I don't think so because that way I could communicate with English speaking people and I could communicate with Hispanics because that way you don't have troubles understanding what I'm saying."

Because of the claims by Daniels (1995) that students would fear obtaining a native accent, the interviewers asked both groups, "Do you feel any fear towards obtaining a native accent?" All of the ELL students said no, many expressing surprise or shock that we would think to ask such a question. One student responded, "No, totally not because I've studied somewhere and gaining that accent doesn't really do anything bad to me. It's like more beneficial." While all of the bilinguals also clearly said no, some hesitated or showed some concern about their other language. Maria said, "Um I don't think I have a fear of it… the only thing right now is the problem…that that native accent is really pushing that other language…so as long as I can balance both of them I wouldn't be worried about it."

Overall in terms of attitudes and expectations for native-like accents, both ELLs and bilinguals gave similar answers. Both ELLs and Bilinguals recognized the advantages of having a native accent. The ELLs considered this to be a goal of their language learning and the bilinguals were not willing to forfeit their native accent in order to show more of their cultural identity. Neither group feared obtaining a native accent. While the ELLs found this question strange to even ask, it is important to note that bilinguals did show some hesitation relating to this question and some mentioned concern for maintaining their other language. Table 1 summarizes the findings for attitudes and expectations.

Table 2
ELL and Bilinguals Attitudes and Expectations

ELL Students	Bilinguals
Prefer to sound like native speakers	Prefer to sound like native speakers
No fear towards obtaining native accent	No fear towards native accent as long as other language is maintained
Would take native accent	Would NOT take a foreign accent

Accent and Identity

To explore the way that students viewed the link between accent and identity and to answer the research question, "How do ELLs and bilinguals perceive a link between their accent in the L2 to their identity?" the ELLS were asked, "If you sounded like a native speaker tomorrow, do you think some of your cultural identity would be lost?" All students said no. Some of the responses are included below.

> **Min:** Umm… I think no.
> **Jing:** Nope. (I: No, why?) Yeah, it's umm I like my culture and then I just want if possible I wanna pull them together, combine them.
> **Gang:** I don't think so (I: Can you explain?) Because it is just the way you talk and not, like cultural identity is just something in your mind but your talk is something in your mouth, so it's not very influential.
> **Xia:** I don't think so actually. Although my accent has been changed to local, but my face and my body language can reflect my Chinese cultural background as well.

The bilinguals were asked the converse question, "Do you ever wish that someone would recognize you as having a foreign accent? Why or why not?" Bilinguals shared mixed responses to this question, three students said no, but two said yes. For example, Tran said, "Not really because then they would start assuming random stuff…rather than just going up to me and asking 'Hey, what's your name…?" while Zamir said, "Yeah kinda because it does bring up the fact that they know I am Albanian"

Students were also asked directly, "Do you think that your accent in English reflects your cultural identity?" The ELLs all said no. Yang responded, "I think it's…um…how to say it… the culture influence the accent, but accent not really reflect back to culture." Bilinguals, when asked the very similar question, "Do you think that a person's accent reflects their cultural identity?" again shared mixed responses. Three bilinguals clearly said yes, while two shared somewhat mixed feelings about the topic. For example, while Manny gave a limited recognition of the link by saying, "It gives you an indication of where they are from, and it gives you an idea of how long they have been staying in the U.S…" Adanya gave a clear yes saying, "I'd say [my accent] represents who I am, but I dunno since I don't have it as thick anymore maybe I am kinda losing-I'm definitely losing it so maybe my identity is going away with it…"

Overall, when looking at the link between accent and identity, differences between the two groups emerge. The ELLs perceived their accent as separate from their identity, mentioning their own perception of identity or appearance as more important to identity. The ELLs saw a native accent as facilitating conversation and relationships with Americans, but did not think that a native accent in English would inhibit or deteriorate relationships with their cultural group. The bilinguals recognized much more of a link between their accent and identity. Many of them saw it as one of the indicators of their cultural identity. While bilinguals agreed with the ELLs that a native accent can help connect with Americans, some stated that the native accent can also inhibit their connection with their cultural group. Table 2 summarizes the findings regarding accent and identity.

Table 3

ELL and Bilinguals Views on Accent and Identity

ELL Students	Bilinguals
Perceive their accent as separate from their identity	Give more recognition to their accent as part of identity
Native accent helps connect with Americans	Native accent helps connect with Americans
Native accent DOES NOT inhibit connection with cultural group	Native accent CAN inhibit connection with cultural group

DISCUSSION

Although Daniels (1995) argued that students must fear obtaining a native accent, this is not supported by the results of this study. Neither group expressed fear of a native-like accent. Instead both groups acknowledged the benefits of obtaining a native-like accent. The experience, however, of obtaining a native accent in English did seem to alter the perceptions of many of the bilinguals regarding the link between accent and identity. The bilinguals overall recognized more of a link between accent and identity. Generally, bilinguals did recognize that accent can serve as an indicator of identity, but many pointed to other features that also contributed to their identity, such as knowing the second language, group affiliations on campus, choices of material goods such as clothing and cars, meeting society's expectations for the cultural group (comparing themselves to stereotypes of the cultural group), and appearance.

Bilinguals also realized that having a native accent in English could impede their ability to connect with their cultural group. While the bilinguals all valued knowing two languages, some missed being able to make immediate connections with their cultural group (for example, when meeting a new person on campus the bilingual might recognize due to accent that the other person is from the same country, but the other person is not able to make the same recognition).

While these results do begin to shed light on the differences between these two populations, there are limitations to this study that must be recognized. First, this study was only able to interview a total of 13 students, a small sample. Additionally, the ELL group was largely from a Chinese background. While this is representative of the international student population at the university,

these students may not be representative of the whole ELL population. Further, the ELL students were all students that had already chosen to study in an English speaking country. This may have affected the way that students saw the benefits of a native accent.

In the future, it would be useful to replicate this study on a much larger scale. More participants should be interviewed from more language learning backgrounds (e.g. heritage learners). Further, it would be useful to interview different generations. With technology changing opportunities for language study and the role of English as a lingua franca, older and younger generations may hold different views regarding the usefulness of a native-like accent.

ABOUT THE AUTHORS

Shannon McCrocklin is an assistant professor at the University of Texas-Pan American. She earned her Ph.D. in Applied Linguistics and Technology at Iowa State University in Ames, IA. She holds an M.A. in Teaching English as a Second Language from the University of Illinois at Urbana-Champaign where she developed an interest in pronunciation teaching and applied phonetics and phonology. Shannon has taught English pronunciation to undergraduate and graduate students as well as to international faculty at Iowa State and the University of Illinois. Her research focuses on improving pronunciation training for students and CAPT (Computer-Assisted Pronunciation Teaching). She has presented at CALICO, NCTE, PSLLT, and AAAL.
Stephanie Link is a doctoral student in the Applied Linguistics and Technology program at Iowa State University. Her research interests are in the use of developmental SLA theory and systemic functional approaches to the design of CALL technologies and automated writing evaluation for assessing students' writing development. Currently, she is a research assistant for the graduate college, working on discourse analysis and tool validation. She is also an instructor of Introduction to Linguistics and Preparing Publishable Thesis Chapters. Her secondary research interests involve pronunciation acquisition, specifically the development and construction of identity through L2 accent and the influence of native versus nonnative speaker pronunciation instructors on students' comprehensibility/accentedness and material development.

REFERENCES

Andreasson, A. M. (1994). Norm as a pedagogical paradigm. *World Englishes*, *13*, 195-409.

Auer, P. (Ed.). (2002) *Code-switching in Conversation: Language, interaction, and identity*. New York: Taylor & Francis

Bresnahan, M. J., Ohash, R., Nebashi, R., Liu, W. Y., & Shearman, S. M. (2002). Attitudinal and affective response toward accented English. *Communication Abstracts*, *25*(5), 591-750.

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching Pronunciation: A course book and reference guide.* (2nd ed.). New York: Cambridge University Press.

Dalton, C., & Seidlhofer, B. (1994). *Pronunciation*. Oxford: Oxford University Press.

Daniels, H. (1995). Psycholinguistic, psycho-affective and procedural Factors in the acquisition of authentic L2 phonology. *Speak Out!, 15*, 3-10.

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

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

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. *Language and Speech, 38,* 25–56.

Haslam, A. S. (2001). *Psychology in organizations: The social identity approach.* London: Sage Publications Ltd.

Johnstone, B. & Kiesling, S.F. (2008). Indexicality and experience: Exploring the /aw/ monothongization in Pittsburg. *Journal of Sociolinguistics*, *12*, 5-33.

Levon, E. (2006). Mosaic identity and style: Phonological variation among Reform American Jews. *Journal of Sociolinguistics*. 10(2), 181-204.

Marx, N. (2002). Never quite a 'native speaker': Accent and identity in the L2 – and the L1. *Canadian Modern Language Review*, *59*, 264-281.

Merriam, S. B. (2002). *Qualitative research in practice: Examples for discussion and analysis*. San Francisco: Jossey-Bass.

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

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

Norton, B. (1997). Language, identity, and ownership of English. *TESOL Quarterly, 31* (3), 409-429.

Ochs, E. (1993). Constructing social identity: A language socialization perspective. *Research on Language and Social Interaction*, *26*(3), 287-306.

Rindal, U. (2010). Constructing identity with the L2: Pronunciation and attitudes among Norwegian Learners of English. *Journal of Sociolinguistics.* 14, 240-261.

Setter, J. & Jenkins, J. (2005). State-of-the-art review article. *Language Teaching*, *38*(1), 1-17.

Shin, S. J. (2010). 'What about me? I'm not like Chinese but I'm not like American.' Heritage language learning and identity of mixed heritage adults. *The Journal of Language, Identity, and Education, 9*(3), 203-219.

Sobkowiak, W. (2005). Why not LFC?. In K. Dziubalska-Kołaczyk & J. Przedlecka (Eds.), *English Pronunciation Models: A Changing Scene* (pp. 131–49). Frankfurt: Lang.

Yamaguchi, C. (2002). Towards international English in EFL classrooms in Japan. *The Internet TESL Journal*, 8. [http://iteslj.org/Articles/Yamaguchi-Language.html. Last accessed 19th January 2011]

Muller Levis, G., & Levis, J. (2014). Using introductions to improve initial intelligibility. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 145-150). Ames, IA: Iowa State University.

Teaching Tip:
Using introductions to improve initial intelligibility

Greta Muller Levis – Iowa State University
John Levis – Iowa State University

Early instruction in speaking any language includes pre-packaged chunks of functionally important language such as introductions and leave-takings. Such functional uses of language are critically important in early language use (e.g., Weinert, 1995) and provide a way for beginners to communicate in highly constrained situations (Wray, 2000). In addition, learners with higher proficiency can use the opportunities provided by functional routines to gain access to social interactions and thus, access to greater opportunities for learning the target language. However, the success of functional routines is often dependent on pronunciation that matches the expectations of interlocutors.

Goals of Teaching Tip

1. To demonstrate the importance of pronouncing names in a listener-friendly way.
2. To describe how pronunciation impacts small talk routines.
3. To show how greeting routines operate with 2 and 3 people together.

This teaching tip demonstrates how introductions in English can be used to introduce communicatively-oriented pronunciation from the very beginning of instruction, whether that is in a class for beginners or as an ice-breaker to start an oral communication class at higher proficiencies. The teaching tip includes three elements that promote pronunciation, pragmatics and oral communication. The three elements can be used individually, together or in sequence: Saying personal and place names clearly, making small talk, and introducing one person to another. Although the tip talks about introductions in English, we also give suggestions for ways that the principles behind the tip may apply to other languages.

This teaching tip is originally based on Wayne Dickerson's exercise, *Stress on Names*. The original exercise focused on the pronunciation of personal and place names in English. Names follow the default prosodic pattern in English in that the focus always falls on the last part of the name. Like other English sentences, the last part of the name is overwhelmingly on the stressed syllable of the last lexical word (or content word) in the phrase. Crystal (1969) found that the last content word of the phrase carries focus about 90% of the time in normal spoken English. For example, in (1), the stressed syllables of *Nelson* and *Iowa* are the default location in each phrase for focus. Focus in English is marked by pitch and length changes.

(1) My name's Jim Nelson. I live in Ames, Iowa.

Proper use of focus (also called primary stress, sentence stress, nuclear stress, prominence and tonic by other writers) is very important for how English speakers understand speech. Hahn

(2004) found that English-speaking listeners understood lectures better when words were spoken with the correct focus. In addition, listeners evaluated speakers more positively when the speakers used focus correctly.

In one sense, names are no different from any other use of focus. However, in another sense, names are unusual. Names carry a higher communicative burden because of their role in greeting routines. We start conversations with new people by providing our names. Because names are spoken in particular ways, they can help in a successful beginning of an interaction, or they can be a reason that an interaction struggles. If names are spoken in expected ways, they are easier to hear. If they are spoken in unexpected ways, possible conversations can stumble from the start.

Personal names in particular carry a very large communicative load. People know how to say their names in familiar ways, and if the way they say their names does not sound right to English-speaking listeners, those listeners will not hear them accurately. And as a result, the listener may ask the speaker to repeat the name, which is likely to be pronounced in the same way the L2 speaker is used to saying it but not in the same way the native listener is used to hearing it. This is not a recipe for success.

Exercise 1 – Self introduction

The first part of this teaching tip involves learning to say names in the way that English speakers are used to hearing them. An easy way to do this is in using a chain exercise in which people introduce themselves to the next person, as below. This kind of exercise serves to make student familiar with the names of their classmates and to help the teacher know what kind of functional pronunciation the learners have with this basic task.

Exercise 1: Introduce yourself to the person next to you. Then that person introduces herself/himself to the next person, and so forth.

Person 1: Hi, I'm _____ _____.
Person 2: Hi, my name's _____ _____.

(Continue with Person 2 introducing self to Person 3, and so on).

Exercise 2 – Saying names "English style"

After this exercise, it is time to draw attention to the way that names are said in English by pointing out the prosodic patterns expected by English listeners. Learning to say names in this way is what Olle Kjellin, a Swedish pronunciation expert, calls "listener-friendly pronunciation". The next exercise involves identifying the part of the name that gets the greatest emphasis. The way such emphasis is signaled is not pointed out ahead of time, as students usually pick up on which words are more prominent. It is possible to then do repetition practice with names if time allows.

Exercise 2. Listen and identify which part of the name gets the greatest emphasis.

Richard Smith	Jason Nelson	Professor Delling	Doctor James Fallon
Anne Marie	Emily Carter	Mrs. Smithson	Reverend Jensen
Des Moines	San Francisco	Hong Kong, China	Boston, Massachusetts

Names in English are emphasized on the _____.

The take-away lesson from this exercise is that names in English are spoken with a particular emphasis pattern, with emphasis on the last part of the name. This pattern is true for personal and place names. Our experience is that students do not have trouble saying this pattern for place names, since they do not have a personal commitment to how these names are emphasized. However, they often do have trouble saying their own names according to the English pattern. Their names packaged in English prosodic wrapping can sound strange, funny or wrong. And personal names are part of personal identity. Making changes to the way one's name is said, even if it helps listeners to understand, can be difficult. Saying names in a new way is not an issue of cultural imperialism. It simply is a recognition that L2-L1 interactions are inherently unequal, and that L1 listeners may need help in processing an unfamiliar name. Names spoken with an unexpected emphasis pattern are harder for English speakers to hear and process. This is true for all names, even those that are familiar. If the name is unfamiliar and is spoken with an unexpected emphasis pattern, it can be almost impossible for English listeners to hear clearly.

Exercise 3 – Small Talk

Names are important, but they are not the only part of successful introductions. Exchanging names leads to small talk, that is, talking about routine and non-threatening topics that lubricate conversational opportunities. These kinds of functional routines also have prosodic features that are critical to their success in English. The exercise below builds on the use of self-introductions and includes a typical small-talk routine. In this case, the small talk involves the question "Where are you from?" which is asked by Person 1 and then repeated by Person 2.

The first asking follows the default pattern for focus in English, with emphasis on the last word, FROM. However, the second asking reflects a new pattern used in repeated questions. In this second question, emphasis shifts to YOU, reflecting the new referent of the pronoun. The pattern is common for questions of this sort, as in the well-known phatic exchange *How ARE you? Fine. How are YOU?* The return question doesn't have to be exactly the same, but the shift of emphasis to YOU remains characteristic. For example, in *What are you DOing? Not much. How about YOU?* the repeated question is a general abbreviated form of the first question, yet still has the same pattern.

Exercise 3. Self-introductions – Introduce yourself to someone else. Use the expected emphasis patterns on the final part of your name in your introduction. Then switch and introduce yourself to another person.

Person 1: Hi, my name's _____ _____.
Person 2: Hi, I'm _____ _____. Where are you FROM?
Person 1: I'm from _____, _____. Where are YOU from?
Person 2: I'm from _____, _____.

Self-introduction and small talk in English often include repeated questions (How ARE you?, Fine. How are YOU?). Such questions have a particular emphasis pattern that shifts from the first asking to the second. The second one always has the emphasis on YOU, marking the shifting referent of the word "you".

An advantage of using repeated questions is that students seem to intuitively understand the communicative use of the focus pattern and the repetition. The pronunciation is critical to the language function, and the function itself is useful in many different contexts.

Exercise 4 – Peer introductions

The final exercise in this teaching tip is one we typically do later, recycling the practice in a slightly more complicated way. This is an exercise that we usually reserve for intermediate or more advanced students. This further practice can take place with peer introductions, in which students work in pairs to talk to introduce themselves to each other and to ask small talk questions before introducing each other to the rest of the class. Beyond this, self and peer introduction can be done in groups of 3, and can involve introducing one person to another. These kinds of three-way introductions highlight other grammatical and pragmatic issues, especially the use of "This is" rather than "S/he is" and the use of body language to introduce one person to another.

Exercise 4. Introducing others – In groups of 3, Person 1 introduces Person 2 to Person 3. Follow the pattern and use the expected emphasis patterns.

Person 1: _____, (1) this is _____ _____.
(first name 1) (first name 2) (last name 2)

S/he's from _____, _____.

(2) _____, _____. _____, _____.
(first name 2) (first name 1) (first name 1) (first name 2)

Person 2: Hi, __(first name)___.
(3) It's nice to MEET you.

Person 3: Hi, __(first name)___. (4) Nice to meet you, TOO.

Exercise filled in with possible names.

Person 1: Olga, this is Peter Schmitt.

He's from Geneva, Switzerland.

Peter, Olga. Olga, Peter.

Person 2: Hi, Peter.
It's nice to MEET you.

Person 3: Hi, Olga. Nice to meet you, TOO.

This kind of introductory routine is unusual in ESL materials, although it is not uncommon in contexts where people are meeting for the first time (as in "cocktail party" type introductions). There are four elements in this short dialogue that bear mentioning. (1) highlights that introducing someone involves the use of the somewhat impersonal sounding "This is…" rather than the more intuitive "She is.." or "He is…". Our experience is that students use "S/he is…" and sound strange. In (2), the repeated introductions to each other offer a chance to remember the names, but also are accomplished with hand motions, in which names are said with a hand signal and a quick visual focus. The hand motion and visual focus then move to the other person as the other name is said. In (3), another functional routine is used for more attention to focus, with "It's nice to MEET you". A variation in the change of focus to you shows up in (4), in which the use of TOO receives focus. TOO often shows up in its own phrase when used in final position, and in that position, receives the focus.

Final thoughts about names and introductions

Greetings and introductions are critical in all languages and have culture-specific communicative patterns that involve pragmatics, language content, and pronunciation features. They are essential parts of phatic communion in that they allow people to begin interactions with others and thus open to door to further interaction. To be able to successfully introduce oneself, speakers must sound right, or at least be within a range of acceptability so that interactions do not fail before they even start.

One of the biggest obstacles to success is how personal names are said. Names are as close as our skin, and saying one's personal name in an unfamiliar way can be disconcerting or even offensive to the speaker. Unfortunately, a target language speaker may have to hear a name in such a way in order to understand it. No amount of trying to enlighten a target language speaker will help them hear, especially in the short run. And introductions are decidedly short-run types of interactions.

Pronunciation features are language specific. In English, they include:

a. Emphasis, or focus on the last part of the name, no matter how long the name, and no matter whether the name is personal or a place.
b. Shifting emphasis to YOU on questions returned in much the same form, e.g.,

A: Where are you FROM?
B: _____. Where are YOU from?
A: Cool. I'm from _____.

In other languages, the pronunciation and content patterns will differ but will likely play just as important of a role. Teachers should not expect that most learners will pick up these patterns without help. The costs of failure are high enough that explicit practice on these kinds of interactions will pay off many times over. Understanding how to say one's name so listeners understand it builds success in first interactions, makes clear the importance of pronunciation from the beginning of instruction, and helps learners understand the cultural aspects of typical greeting routines.

ABOUT THE AUTHORS

Greta Muller Levis is a Senior Lecturer in TESL and Applied Linguistics at Iowa State University, where she also is the adviser for undergraduate students in linguistics. She has taught international teaching assistants for over 20 years. Her interests include the teaching of pronunciation, effective teaching practices and teaching introductory linguistics. She has presented papers at conferences in the United States, Canada and Taiwan.

John M. Levis is Professor of TESL and Applied Linguistics at Iowa State University. He teaches a variety of courses including Methods of Teaching Pronunciation, Listening and Speaking, Linguistics, and Sociolinguistics. He is the organizer of the Pronunciation in Second Language Learning and Teaching conference and has published in a variety of journals, including *TESOL Quarterly, Applied Linguistics, World Englishes, ELT Journal, Annual Review of Applied Linguistics,* and *System*. He is a co-editor of *Soci*al *Dynamics in Second Language Accent* (DeGruyter Mouton, 2014) and of the *Handbook of English Pronunciation* (Wiley Blackwell, 2015).

REFERENCES

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201-223.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, *16*(2), 180-205.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, *21*(4), 463-489.

Ruellot, V. (2014). Introducing French nasal vowels at the beginner level: A demystification. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 151-55). Ames, IA: Iowa State University.

TEACHING TIP
INTRODUCING FRENCH NASAL VOWELS AT THE BEGINNER LEVEL: A DEMYSTIFICATION

Viviane Ruellot, Western Michigan University

Although they make for only a fraction of the French phonological repertoire, nasal vowels represent a notorious challenge in the acquisition of French pronunciation by second language (L2) learners (Dansereau, 1990; Harlow and Muyskens, 1994). As nasal vowels are in phonemic contrast both with oral vowels (*bain/baie, banc/bas, bon/beau*)[10] and with one another (*bain/banc/bon*), their inaccurate perception and production can significantly impact intelligibility. Therefore, an introduction to – and the practice of – these sounds in French pronunciation is warranted. The results of the informal polls I have conducted as an instructor of French over the years seem to indicate that one initial obstacle to feeling comfortable with French nasal vowels is linked to the notion held by some learners that these sounds represent an entirely foreign concept, that they simply do not exist in the native English sound repertoire. Consequently, these learners are unable to relate French nasal vowels to familiar sounds and find it very difficult to process them. The present teaching tip focuses on the *introduction* of French nasal vowels, i.e., the stage preceding any practice with the perception and production of these sounds. Its goals are to "demystify" French nasal vowels by demonstrating to students that vowel nasality exists in English (Ruhlen, 1973) and by guiding learners to pair unfamiliar French sounds with familiar English sounds (Levac, 1991), so that they can resort to these referents when processing French nasal vowels during perception and production practice.

The first step of this approach concerns the existence of vowel nasality in English. It is based on pronunciation information featured in two French textbooks co-authored by Valdman (*Chez Nous*, 2010, and *Points de Départ*, 2013): "Both English and French have nasal vowels. In English, any vowel followed by a nasal consonant is automatically nasalized, as in *man*, *pen*, *song*." While this concept – repeated *verbatim* in both texts – is mentioned, it is not further developed. Nor is it followed by activities for practice. The second step of this approach is informed by Levac (1991), who recommends pairing unfamiliar sounds (here, French nasal vowels) with familiar ones (i.e., English nasalized vowels), so as to provide learners with "starting points for 'anchoring' a French vowel phoneme" (p. 216).

Background

The French phonological repertoire consists of 21 consonants and semi-consonants, 12 oral vowels, and 3 nasal vowels.[11] While they represent a small portion of the French sound system,

[10] *Bain* /bɛ̃/, 'bath' vs. *baie* /bɛ/, 'bay'; *banc* /bɑ̃/, 'bench' vs. *bas* /ba/, 'low'; *bon,* /bɔ̃/, 'good' vs. *beau* /bo/, 'beautiful.'

[11] The fourth nasal vowel /œ̃/ was excluded from this list because research shows that it is gradually disappearing from French speech, with more and more native speakers substituting /ɛ̃/ for /œ̃/ (Hansen, 1998).

L2 learners typically find it difficult to discriminate between /ɛ̃/ (as in *bain*), /ɑ̃/ (as in *banc*), and /ɔ̃/ (as in *bon*). This difficulty may well be linked to the fact that the three sounds are acoustically closer to each other than are their oral counterparts (Lindblom, 1975) and, consequently, they are not as readily distinguishable from one another. Possibly added to this difficulty is unawareness – in some learners – of the existence of vowel nasality in English. Indeed, when an English vowel is adjacent to a nasal consonant (as in 'fan'), it is automatically nasalized. Because it is due to physiological constraints linked to the anticipated pronunciation of the nasal consonant, vowel nasality is not phonemic in English. For instance, the word 'fan' is phonemically /fæn/ but it can be pronounced [fæ̃n], with no impact on its meaning. For this reason – and also because a smaller quantity of air passes through the nose during nasalization – vowel nasality is only partial in English (Ruhlen, 1973). On the contrary, in French, it is full and intrinsic (i.e., not as a consequence of the phonetic environment) and phonemic (e.g., *banc* /bɑ̃/, 'bench' vs. *bas* /ba/, 'low').

Exercise 1 – English Nasalized Vowels

While the preceding explanation is probably considered wholly pertinent by most phonologists and phoneticians, a different – less technical – approach is preferable as the first step of this teaching tip. Indeed, the demystification of vowel nasality and students' awareness of its existence in English can simply begin with an exercise that focuses on the vibration caused by the air passing through the nose during the production of a partially nasalized English vowel. Students are invited to place their thumb and index around their nose and repeatedly pronounce the pairs of English words below, the first one containing an oral vowel, and the second one a partially nasalized one. Alternating between lack of vibration and presence thereof will help students realize that English does have nasalized vowels, so that the notion that vowel nasality is a foreign concept can begin to fade from the students' mind.

Exercice 1: Place your thumb and index around your nose and pronounce the following pairs of words. Feel the vibration as you pronounce the second word.

'bad' – 'band'
'log' – 'long'
'thick' – 'think'

The concept of vowel nasality should be regularly re-visited for reinforcement through practice of additional word pairs:

'cap' - 'camp' / 'cat' - 'can't' / 'dap' - 'damp' / 'lad' - 'land' / 'lap' - 'lamp' / 'tack' - 'tank'
'bod' - 'bond' / 'chop' - 'chomp' / 'cop' - 'comp' / 'pod' - 'pond'
'chip' - 'chimp' / 'pig' - 'ping' / 'sick' - 'sink'
'bet' - 'bent' / 'fed' - 'fend' / 'led' - 'lend' / 'let' - 'lent' / 'med' - 'mend'
'truck' - 'trunk'

Exercise 2 – French and English Nasal Vowels

The second step of this teaching tip consists of comparing French nasal vowels and partially nasalized English vowels, both in terms of perception and articulation. This step, designed to further ease students into the concept of vowel nasality, will allow them to relate unfamiliar sounds (i.e., French nasal vowels) to familiar ones (i.e., English partially nasalized vowels). Each of the three French nasal vowels is associated to a partially nasalized English vowel through French–English word pairs. These words were selected for their imperfect but sufficiently close visual and acoustic resemblance.

A. /ɛ̃/: French *fin* and English 'fan'

The sound /ɛ̃/ is presented with the French-English word pair *fin* ('end') and 'fan.'[12] Students are informed that the vowels in *fin* and 'fan' are very close in terms of perception and articulation, although for the beginning of [æ̃] in 'fan,' the back of the tongue may be slightly higher and more forward than it is for [ɛ̃] in *fin*. However, this difference seems to be too imperceptible to significantly affect how useful it is for students to liken the vowel in *fin* to the vowel in 'fan.' Students are invited to first pronounce the word 'fan' without the /n/ and the French word *fin* repeatedly, paying attention to how closely both vowels sound and how similar the configuration of their articulators is as they pronounce both vowels.

Exercise 2A: First pronounce the English word 'fan' without the final 'n' sound and then, the French word *fin*. Repeat the word pair several times. Note how similar both vowels sound, and try to visualize the shape of your tongue (its back is high and its tip is low) as you pronounce both vowels.

B. /ɑ̃/: French *langue* and English 'long'

For a reference to the sound /ɑ̃/, English 'long'[13] is likened to French *langue* ('tongue'). Again, the back of the tongue at the beginning of the articulation of the partially nasalized vowel in 'long' starts in a slightly higher position than it does for the /ɑ̃/ in *langue*. However, this difference does not significantly affect the usefulness for students of likening the two vowels and creating a reference to anchor unfamiliar French [ɑ̃]. As in A., students pronounce 'long' and *langue* repeatedly, focusing on the similarity between the two vowels in terms of perception and articulation.

Exercise 2B: Pronounce the English word 'long' and the French word *langue* repeatedly. Note how similar both vowels sound, and try to visualize the position of your tongue (its back is lower than for *fin*) as you pronounce both vowels.

C. /ɔ̃/: French *on* and 'honhonhon!'

Finally, the sound /ɔ̃/ as in *on* (indefinite subject pronoun, singular) is referred to the repeated sound used to parody French laughter 'honhonhon,' when articulated like French /ɔ̃/, i.e., with a

[12] The pronunciation of the word 'fan' can be heard on the MacMillan Dictionary website at http://www.macmillandictionary.com/pronunciation/american/fan.

[13] To hear the pronunciation of the word 'long' as referred above, go to the Oxford Advanced Learner's Dictionary at http://oald8.oxfordlearnersdictionaries.com/dictionary/long_1, and click the button to the right of the phonetic transcription [lɑ:ŋ].

very closed mouth and the back of the tongue raised high. Although this comparison is likely to reinforce stereotypes, it has the virtue of including an expression that is familiar and amusing, both qualities that are likely to ease students' processing of this French nasal vowel.

Exercise 2C: Pronounce the French word *on* as you pronounce the vowel in the stereotypical French laugh 'honhonhon.' Make sure your mouth is very closed and the back of your tongue so high that it almost touches the roof of your throat.

CONCLUSION

Some learners, especially those who do not have a natural ear for sounds, need strategies to help them process the unfamiliar sounds of a foreign language before they can start practicing their perception and production of those sounds. This teaching tip provides such learners with strategies for tackling French nasal vowels, by familiarizing them with the concept of vowel nasality and by guiding them to use familiar sounds from their native language as referents for unfamiliar foreign sounds until the latter have become familiar. The central idea beyond this approach may not be readily applicable to all of the unfamiliar sounds of a second language, but for those sounds where some semblance of an equivalent exists, this strategy will be effective for L2 acquisition.

ABOUT THE AUTHOR

Viviane Ruellot is an Associate Professor of French at Western Michigan University, where she teaches French, French linguistics and applied linguistics. She serves as French section head, coordinator of beginning and intermediate French courses, and liaison between Western Michigan University and Kalamazoo high schools. Her research focuses on the pedagogy and acquisition of French pronunciation by non-native speakers. She studies how feedback may help learners bridge the gap between perception and production and improve their pronunciation. She is also interested in the stages of pronunciation acquisition and the history of French pronunciation teaching.

REFERENCES

Dansereau, D. (1990). *Savoir dire : Cours de phonétique et de prononciation*. Lexington, MA: D.C. Heath.

Hansen, A. (1998). *Les voyelles nasales du français parisien moderne. Aspects linguistiques, sociolinguistiques et perceptuels des changements en cours*. Copenhague, Museum Tusculanum Press.

Harlow, L. and Muyskens, J. (1994). Priorities for intermediate-level language instruction. *Modern Language Journal, 76*, 491-501.

Levac, P. (1991). Teaching languages communicatively: An examination of the presentation of phonology to first-year students of French. In J. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics 1991* (pp. 215-222). Washington, DC: Georgetown University Press.

Lindblom, B. (1975). *Experiments in sound structure*. Plenary address, 8th ICPS, Leeds.

Ruhlen, M. (1973). Nasal vowels. *Working Papers on Language Universals, 12*, 1-36.

Scullen, M.E., C. Pons, and A. Valdman. (2013). *Points de Départ* (*2nd ed.*). Upper Saddle River, NJ : Prentice Hall.

Valdman, A., C. Pons, and M.E. Scullen. (2010). *Chez Nous. Branché sur le Monde Francophone* (*4th ed.*). Upper Saddle River, NJ : Prentice Hall.

Saalfeld, A. (2014). Flipping the phonetics classroom: A practical guide. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 156-171). Ames, IA: Iowa State University.

FLIPPING THE PHONETICS CLASSROOM: A PRACTICAL GUIDE

Anita Saalfeld, University of Nebraska at Omaha

INTRODUCTION

Although multiple studies (Derwing, Munro, & Wiebe, 1997; 1998; Gónzalez-Bueno, 1997; Hahn, 2004; Kissling, 2013; Lord, 2005; 2008; Miller, 2012; Sturm, 2013a; 2013b) investigating classroom acquisition of second language (L2) phonology have indicated that learners can improve their pronunciation with instruction and/or practice, Saalfeld's (2011) study revealed no improvement in pronunciation for learners enrolled in a Spanish phonetics course for Spanish stress placement, and subsequent unpublished data analysis revealed that learners did not improve in any category after a semester of instruction. A key finding of Saalfeld (2011) was that students enrolled in the phonetics course exhibited statistically significantly better pronunciation than students in the control group at the outset of the study, indicating that students who would have benefitted most from a phonetics course elected not to take it. Nevertheless, in most phonetic categories, there remained substantial room for improvement by students enrolled in the phonetics course, which raised the following questions: Why did students not make significant gains in their pronunciation? What changes could be made to the course to promote improvement of pronunciation?

As part of a redesign of the math program at the University of Nebraska at Omaha, the university purchased a membership to the National Center for Academic Transformation (NCAT), which included a number of various on-campus workshops on course redesign. In addition, work samples from other course redesign projects were freely available on the NCAT's website. The goal of these redesign projects is to use existing technology to improve learning outcomes, and there are a number of models that can be used to achieve this goal. For the current project, the 'flipped' classroom model described by the NCAT was selected due to its emphasis on classroom practice time. Essentially, a flipped classroom moves part or all of the instructional components outside of regular class time, and class time is spent on student-centered tasks such as discussions, practice, direct application of concepts, and other forms of active learning, rather than on teacher-centered passive activities such as lectures. This seemed appealing for a phonetics course, given the emphasis on practice to improve pronunciation. In this particular flipped model, in addition to time devoted to pronunciation practice in class, there was also time designated each week for students to work on homework assignments during class time so that they could receive help as problems arose.

The main question guiding the implementation of the flipped phonetics course was: What changes needed to be made to promote student learning and improvement of pronunciation? The clear answer was that students needed more practice time. There were two obvious ways to achieve this: first, designate more in-class time for pronunciation practice, and second, assign more recorded homework assignments. However, these

options presented several challenges. In the original iteration of the course, although there was time dedicated to pronunciation practice in nearly every class period, it was impossible to monitor students to ensure that they stayed on task. Given this issue, increased practice time without other modifications to the course delivery format might have had negligible effects. For the recorded homework assignments, the original iteration of the class included a recording every other week. Since this is only one course of three that constitute a typical semester teaching load, adding more recorded assignments needing evaluation would have created an unreasonable amount of work for the instructor. The key question then became: How could additional practice be added to the course without increasing instructor workload?

Implementation

1. Digitize written homework assignments. In the first iteration of the course, there was a bimodal grade distribution on written homework assignments and quizzes, with about half of the class at an A average, and the other half at a low C average. The students in the A-range thought that the course moved far too slowly, and the students in the C-range thought the course moved far too quickly. The C-range students needed more practice with all of the concepts, but the A-range students would have been bored and uncooperative if additional class time had been spent on something that they did not struggle with.

Migrating all of the written assignments to a digital format using the university's course management system (Blackboard) provided a solution to this problem, as it allowed students to repeat the assignments in order to master the concepts presented in class. Students who immediately understood the concepts were able to do the assignment once and move on to the next topic, but students who struggled were able to repeat the assignment and determine where they were having problems on an individual basis. This was a key advantage of digitization. With a traditional model, it would be logistically impossible to allow students to repeat assignments due to the detailed, tedious nature of grading transcriptions.

In order to maximize learning, I implemented the following conditions on the digitized homework assignments.

a. Students were allowed to repeat assignments, but only a limited number of times[1]. Since students weren't able to keep doing the assignment until they figured out the correct answers by chance, they could not simply guess their way to the correct answers without learning.
b. Students could immediately see the answers they submitted for all questions, but they were not able to see the correct answers until after the deadline for an assignment had passed. This prevented them from copying and pasting the correct answers without understanding what they were doing wrong.
c. In the first iteration of the flipped course, I implemented a gated assignment release feature, so that students could only advance to the next homework if they received a grade of 70 or above on the previous one. As a matter of pedagogy and practicality, I discontinued this practice[2]. Pedagogically, it was unnecessary to require students to have mastered one concept prior to moving on to the next one, as they continued to practice that concept, together with a new one, on the next homework assignment.

> Practically, I spent a lot of time making exceptions to deadlines to ensure that students didn't fail the course because of one missing or late homework assignment.

There are a few special considerations necessary for digitized homework assignments in order to maximize student success and minimize frustration with technology.

a. *Give clear directions and a model* (see Figures 1 and 2). This can be easily overlooked when migrating content from a written assignment format to a digital one. With Figure 1 (for all Figures, see Appendix C at the end of the paper), I wanted to encourage productive knowledge of segment descriptions, in which students could produce the description without having a list of options in front of them. Providing specific directions and a model allowed me to use a more active knowledge question format (as opposed to a more passive format such as matching) and still have the question be auto-graded by the course management system. Figure 2 provides an example of an auto-graded paragraph-level transcription question, in which the instructions clearly specify what students need to pay attention to in their transcriptions.
b. *Use matching questions rather than true-false or multiple choice questions* (see Figure 3). If students get a single true-false item wrong, they automatically know the right answer, but they may not know *why* that answer is right and theirs was wrong. For multiple-choice questions, they may be able to guess the right answer without knowing why it's correct, which can be avoided by using a matching question format. In Figure 3, rather than creating 20 multiple choice items asking students to identify the appropriate allophone of /n/, I created a matching question with 20 items and seven answers (the seven allophones of /n/ in Spanish). Students can see their score, but not which individual items they got right or wrong. This way, they're prompted to ask for help on concepts that they don't understand, rather than getting some right answers by guessing, but still not learning anything.

2. Incorporate the use of an acoustic phonetics program into class practice time. To explain the rationale for requiring the use of an acoustic phonetics program, on the second day of class, I played a video of a TED Talk by Dr. Patricia Kuhl (TEDxTalks, 2010; see Appendix A).

The video explains that although humans are born with the ability to perceive sounds in any language, by about 10 months, our brains have learned to ignore sounds that don't occur in our native language(s). Following the video, there was a brief class discussion explaining that because we can't trust our brains to interpret what sounds we hear and produce, we would use an acoustic phonetics computer program that showed pictures of the sounds being produced, so that learners could see the differences in their own productions compared to native speaker productions.

I provided recordings of native speakers, taught learners how to use Praat (Boersma & Weenink, 2013), a free acoustic phonetics program, trained them to analyze various elements of their own speech[3], and taught them how to identify what they needed to do to make their speech more native-like. This strategy made learners responsible for assessing their own pronunciation and figuring out how to improve it, rather than having the instructor as the sole provider of feedback on pronunciation. This approach promoted ongoing learning and improvement, since it teaches learners how to improve their

pronunciation beyond the end of the semester when they don't have access to professor feedback.

3. Increase the number of recorded assignments, but limit grading by selecting only a few to grade over the course of the semester. Although it would be ideal to provide detailed feedback on every assignment, it is not practical and not always a good investment of instructors' time. In order to provide learners with ample practice without burdening the instructor unreasonably, alternate grading procedures can be implemented.

Students were required to submit a recorded assignment on a weekly basis and earned credit based on completion. Additionally, as stated in the syllabus, two assignments per student were graded with detailed comments over the course of the semester, but students had no prior knowledge of which ones would be selected. That method provided sufficient practice opportunities while allowing for detailed feedback and a reasonable workload for the instructor. Students still needed to pay attention to the target sounds in all of the recordings, since they didn't know which were graded. This grading system was adapted from an approach described by a colleague (J. S. Miller, personal communication, November 5, 2013; see Appendix B for a full description).

4. Dedicate the majority of class time to student practice, whether on theoretical material or on pronunciation. In the original implementation of the flipped course, all instruction was delivered outside of class time via videos produced by the instructor using Adobe Presenter. Students indicated that they would have preferred to receive some instruction during the regular class period, so in the next iteration of the course, I provided short explanations of no more than 15 minutes per class (out of a 75-minute class period). The remaining time was devoted to practice of pronunciation and theoretical concepts. The majority of practice activities were converted to an auto-graded digital format so that students could get immediate feedback indicating whether they had understood the concept[4]. During class time, students could not see the correct answers to the practice activities so that they would ask questions during the same class if they did not do well on the practice activities, but the correct answers were made available at the end of each day.

Students were required to complete the homework outside of class time, but there was time designated each week during the class period before the assignment was due for students to ask questions and get clarification about any problems they experienced on the homework. For pronunciation practice, students were required to make recordings in class and use Praat to compare their productions to those of native speakers (provided in the course management system). In order to ensure that students completed the practice activities, I created surveys in the course management system (see Figure 4) that required students to put in information from their recordings, such as formant values for vowels or VOT for stops, for example, and survey completion was part of students' participation grade. This minimized the possibility that students would record themselves without analyzing their own speech production.

5. Schedule the course in a computer lab. The most important element of this course redesign was scheduling the course in a classroom equipped with computers, because the

content for the course was completely digitized and administered through the university's course management system. While most students brought their own laptops to class, not all did. Meeting in a computer lab classroom guaranteed access to the course content during class time.

Special Considerations

There were some challenges and special considerations that complicated the implementation of the flipped phonetics course. In an effort to use resources more efficiently, our campus is currently replacing regular lab computers with virtual clients. This change created a substantial problem during the first iteration of the flipped course, since the virtual clients could not detect analog microphones without having special software installed. Even with special software installed, the virtual clients didn't record sounds at the same level as a traditional machine, so it was necessary to show students how to adjust their audio settings to enable the Microphone Boost feature (and raise the percentage to 100). For this reason, I recommend that instructors inquire ahead of time about what type of computers will be in the lab classroom so that any issues with microphone detection are resolved prior to the beginning of the semester.

One challenge that is likely to occur in Spanish phonetics courses is the growing population of heritage speakers. The heritage speaker population in the current courses consists of active bilinguals whose first language is Spanish, indicating that their phonology is essentially native[5]. Furthermore, research by Au, Knightly, Jun, & Oh (2002), Au, Knightly, Jun, Oh, & Romo (2008), and Knightly, Jun, Oh, & Au (2003) has shown that even childhood overhearers of Spanish enrolled in college-level second-year Spanish courses were able to produce Spanish sounds that were significantly more native-like than those produced by non-native speakers. Therefore, assigning these students recorded homework assignments would create a fundamentally unequal grading scheme, because they would automatically receive a perfect score for 30% of the final course grade by virtue of being heritage speakers. (In this course, 30% of the final course grade was based on digital homework assignments, 30% was based on quizzes, 30% was based on recorded assignments, and 10% was based on class participation.) Equally problematically, this requirement has no benefit for heritage learners, since it is not designed to address their linguistic needs. However, eliminating the recorded homework component for heritage speakers and weighting digital homework assignments and quizzes at 45% each would result in a substantially more difficult course than the 30/30/30/10 weighting, which would be unfair to heritage speakers, both for reasons of relative course difficulty, as well as because it fails to address their specific linguistic needs. There are two possible solutions to this issue. Since one of the main course objectives is improvement of pronunciation, it may make sense to restrict enrollment to only L2 learners. Alternately, it is possible to provide differentiated instruction for both learner populations[6]. For homework assignments, this would involve providing homework assignments of equivalent difficulty focusing on the linguistic needs of heritage language learners as an alternative to the recorded homework assignments. One known issue for heritage Spanish learners is that orthography tends to be challenging, given that a number of sounds are represented by more than one letter[7]. This could be addressed by providing recordings of speakers of dialects that make a phonemic

distinction between the sounds represented by different letters, and asking learners to indicate which word they heard[8]. The larger problem for differentiated instruction is the in-class time dedicated to pronunciation practice. This would require creating separate sets of in-class activities for both learner populations in order to ensure that both benefit from the course. One suggestion for this type of in-class activity would be to have heritage language learners compare recordings of their own speech with that of a speaker of a different dialect of Spanish (I. Velásquez, personal communication, November 18, 2013)[9].

One area that requires further revision is in-class pronunciation practice. Because the assignments were not graded, a number of students treated them as optional, and elected to do other work during the designated pronunciation practice time. Although these students received repeated reminders and lower participation grades when they engaged in this behavior, this was not an effective deterrent. One possible solution is to create graded in-class pronunciation assignments marked as complete or incomplete to ensure that all learners practice pronunciation during the allotted class time.

CONCLUSIONS

The flipped classroom is a valuable model of instruction in phonetics for various reasons. In particular, the ability to repeat homework assignments means that it is possible to implement differentiated instruction in an efficient way. In addition, teaching learners how to evaluate their own pronunciation promotes learner autonomy and ongoing improvement even after the semester has ended. While this model has not led to a 100% student success rate, it has allowed students who would normally struggle and possibly fail the course to get the help and repetition they need in a timely fashion so that they can be successful in the course, both in terms of theory and practice.

ACKNOWLEDGMENTS

I am grateful to Isabel Velásquez (University of Nebraska-Lincoln) for her comments and suggestions for providing differentiated instruction to heritage language learners, and to Jessica Sertling Miller (University of Wisconsin-Eau Claire) for her feedback and insights on earlier drafts of this manuscript. Any errors and omissions are of course my own.
This project was supported by a $1500 curriculum development grant from the University Committee for the Advancement of Teaching at the University of Nebraska at Omaha.

ABOUT THE AUTHOR

Anita Saalfeld is Assistant Professor of Spanish at the University of Nebraska at 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 at Omaha

301C Arts & Sciences Hall
6001 Dodge St.
Omaha, NE 68182
402-554-3298
anitasaalfeld@unomaha.edu

Notes

[1]Students were allowed to repeat an assignment a maximum of five times for the most detailed assignments. The initial homework assignment was the least detailed, and students were allowed to submit it twice. As the content became more detailed, the number of attempts allowed gradually increased to a maximum of five.
[2]I continue to use this gated assignment release structure in an Introduction to Linguistics course, where it is essential that students understand each concept in order to understand the next one.
[3]Some examples include measuring VOT values for voiceless stops, formant values for vowels, noting differences between voiced stops and approximants in a spectrogram, and noting differences between English /ɹ/ and Spanish /ɾ/ and /r/ in a spectrogram.
[4]These practice activities did not count towards learners' course grades; they simply were administered digitally so that learners could repeat them and get immediate feedback.
[5]Although research by Flege (1987) shows that the addition of a second language can affect L1 pronunciation even in late bilinguals, so that it is not equivalent to that of a monolingual native speaker, these differences are detectable only by taking fine-grained acoustic measurements.
[6]As Carreira (2012) notes, it is crucial to differentiate instruction, whether in a heterogenous heritage language (HL) course, or a course made up of heritage language speakers and non-native speakers:

> "Among the most pressing diversity issues surrounding HL teaching are those concerning fairness and accessibility. Given that HL learners with differing proficiency levels are funneled into the same HL class by virtue of constraints on course availability, is it fair and realistic to expect the same of all students in the class? Doesn't this placement outcome virtually doom the less proficient learners to low grades and guarantee high grades to the more proficient learners?
>
> Concerns of this nature often focus on the less proficient learners, since they are the most at risk of failure, both in terms of grade and desertion/truancy. However, there's also cause for concern regarding more proficient learners, as they may be assigned to classes that have little to offer them in the way of new and challenging material. Though these students may end with a good grade, they may be unmotivated to continue their study of Spanish" (p. 102).

[7]The phoneme /s/ is represented with "s", "z" and "ce"/"ci" in all dialects of Latin American Spanish, for example, and /ʝ/ is represented by both "y" and "ll" in most dialects of Spanish.
[8]There are a number of words in Spanish that are homophonous in most dialects of Spanish, but are minimal pairs in a select few. Although "casa" ('house') and "caza" ('hunt' 3rd person singular present indicative) are both pronounced /ˈka.sa/ in all dialects of Latin American Spanish, in Castilian Spanish, the letters "z" and "ce"/"ci" are pronounced /θ/,

and “s” is pronounced /s/, so that “casa” is /ˈka.sa/, and “caza” is /ˈka.θa/. Likewise, although “calló” (‘shut up’ 3rd person singular preterit indicative) and “cayó” (‘fall’ 3rd person singular preterit indicative) are homophonous in most dialects of Spanish, in Andean Spanish, the sequence “ll” is pronounced /ʎ/, while “y” is pronounced /ʝ/, so that “calló” is pronounced /ka. ˈʎo/, and “cayó” is pronounced /ka.ˈʝo/.

[9]Heritage Spanish learners also experience difficulty with placement of accent marks (and connecting spoken word forms to written word forms when multiple prosodic accent patterns are possible), so another possible activity for both in-class and homework assignments could be to provide recordings of minimal pairs or triplets differing only in stress placement and ask them to select the word that they heard. For example, a recording could provide the word ‘calculó’ (/kal.ku.ˈlo/), and learners would need to choose from a set of three words: cálculo (/ˈkal.ku.lo/), calculo (/kal.ˈku.lo/) and calculó (/kal.ku.ˈlo/). These activities, while not strictly related to phonetics and phonology, make use of main concepts discussed in phonetics and phonology (prosodic accent patterns, dialectal variation) to provide instruction specifically designed for heritage language learners.

REFERENCES

Au, T. K., Knightly, L.M., Jun, S., & Oh, J.S. (2002). Overhearing a language during childhood. *Psychologial Science, 13,* 238-243.

Au, T. K., Knightly, L.M., Jun, S., Oh, J.S., & Romo, L. F. (2008). Salvaging a childhood language. *Journal of Memory and Language, 58,* 998-1011.

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.57, retrieved 2 November 2013 from http://www.praat.org/

Carreira, M. M. (2012). Formative assessment in HL teaching: Purposes, procedures, and practices. *Heritage Language Journal, 9,* 100-120.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1997). Pronunciation instruction for "fossilized" learners: Can it help? *Applied Language Learning*, *8*, 217-235.

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

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.

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.

Kissling, E. M. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *Modern Language Journal, 97,* 720-744.

Knightly, L. M., Jun, S., Oh, J. S., & Au, T. K. (2003). Production benefits of childhood overhearing. *Journal of the Acoustical Society of America, 114,* 465-474.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, *88*, 257-267.

Lord, G. (2008). Podcasting communities and second language pronunciation. *Foreign Language Annals, 41,* 364-379.

Miller, J. S. (2012). Teaching pronunciation with phonetics in a beginner French course: Impact of sound perception. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2011. (pp. 109-123). Ames, IA: Iowa State University.

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.

Sturm, J. (2013a). Liaison in L2 French: The effects of instruction. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, Aug. 2013. (pp. 157-166). Ames, IA: Iowa State University.

Sturm, J. L. (2013b). Explicit phonetics instruction in L2 French: A global analysis of improvement. *System, 41,* 3, 654-662.

TEDxTalks (2010, December 13). [Video file]. Dr. Patricia Kuhl. Retrieved from http://www.youtube.com/watch?v=qRRiWg6wYXw

Appendix A
Links to resources

1. http://www.thencat.org/. The National Center for Academic Transformation.
2. http://rishida.net/scripts/pickers/ipa/. IPA character picker.
3. http://ipa.typeit.org/. IPA character picker.
4. http://portfolio.anitasaalfeld.com/samples/blackboard. Additional screen shots of digitized assignments.
5. http://www.youtube.com/watch?v=qRRiWg6wYXw. Dr. Patricia Kuhl's TED Talk on child language acquisition.

Appendix B

Modified pronunciation homework scheme for a flipped phonetics course
(J. S. Miller, personal communication, November 5, 2013)

The grading scheme described in the current article was adapted from one used by Jessica Miller (University of Wisconsin-Eau Claire) in a French phonetics course. In that course, students submitted weekly recorded homework assignments, and the instructor graded one-third of them each week. Miller used an Excel spreadsheet to assign each student a random number, and then sorted them from smallest to largest and divided them into three groups. She then graded one-third of the homework assignments each week during three-week periods, but did not post grades until each three-week period had ended. This way, students did not know whether their assignments were graded during any given week, so that the students whose assignments were graded in the first week would still submit the assignments for the second and third weeks.

Miller has since adapted a completely new assignment structure as a result of flipping her French phonetics classroom. Due to the new instructional model, she has freed up approximately 40 minutes of class time per week for pronunciation practice. This time is used each Friday for conversations. Miller moves from group to group and randomly grades half the class (seven or eight students, in this case) according to rubrics that she has created that evaluate what students should be producing at that point in the semester. She estimates that she spends about 4-5 minutes with each student, giving in-person feedback and direction for resolving pronunciation problems. Students are required to switch partners every week for the conversation time. Miller notes that under this model, students receive five pronunciation grades over the course of the semester, with immediate oral feedback, and that she no longer has to grade oral recordings, which has substantially reduced her grading time.

In addition to the in-class changes, students in Miller's class are required to participate in an hour-long conversation outside of class every other week, and watch a movie every other week. Each Friday when they take a quiz, students turn in a reflection paper prepared at home on either the movie or the conversation (with specific questions related to the lessons and to their own learning), which is part of their quiz grade.

Miller has noted an improvement in students' spontaneous speech production, and that students seem more engaged because the conversation task is more authentic than a recorded homework assignment. Additionally her grading time has been reduced to about an hour and a half per week.

Appendix C - Figures

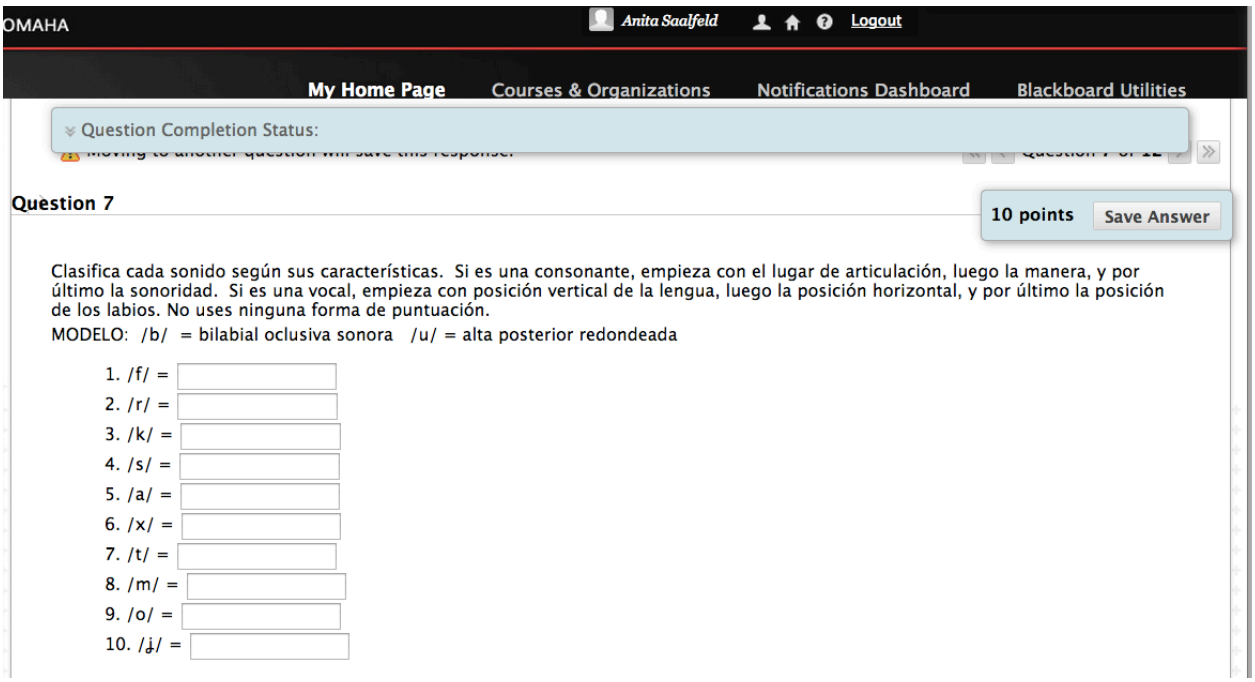


Figure 1. Clear directions and a model. Students were asked to provide a description of the articulation of each segment, using a fill-in-the-blank question type to promote active knowledge.

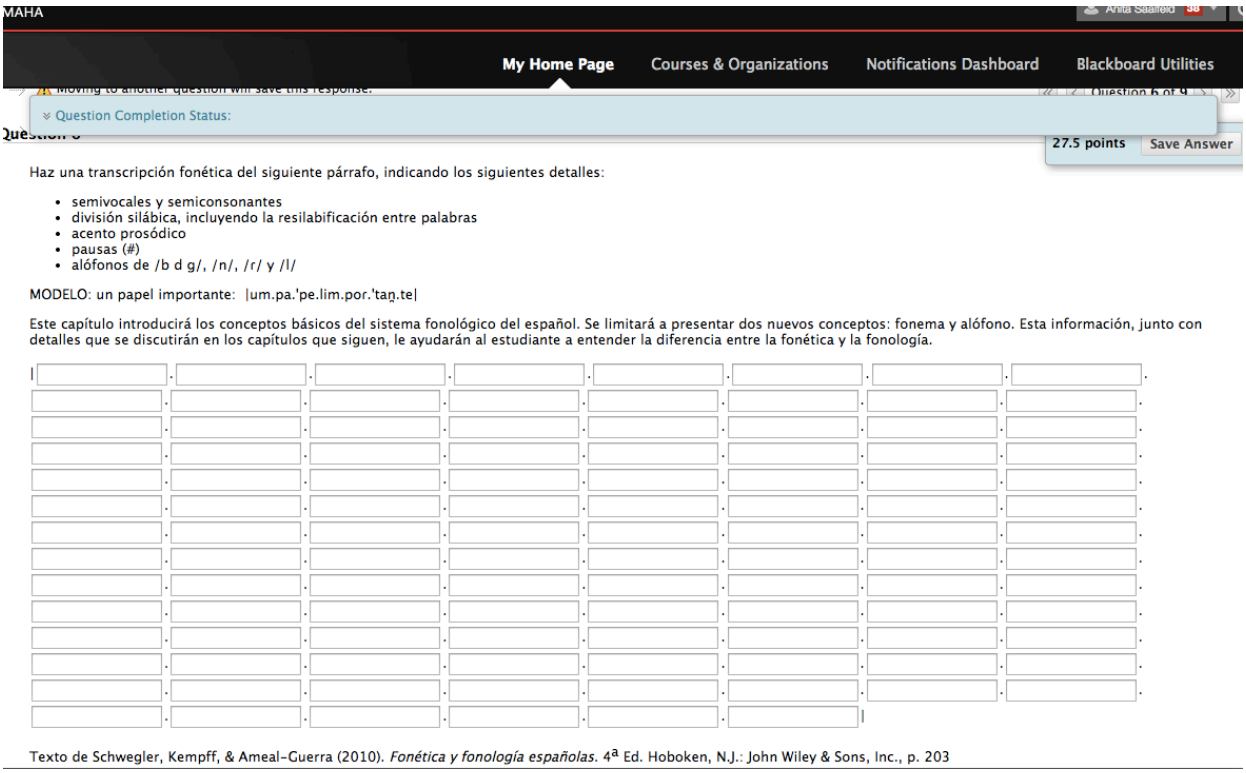


Figure 2. Auto-graded paragraph-level phonetic transcription. Students entered phonetic symbols using an online phonetic character picker (see Appendix A for two options).

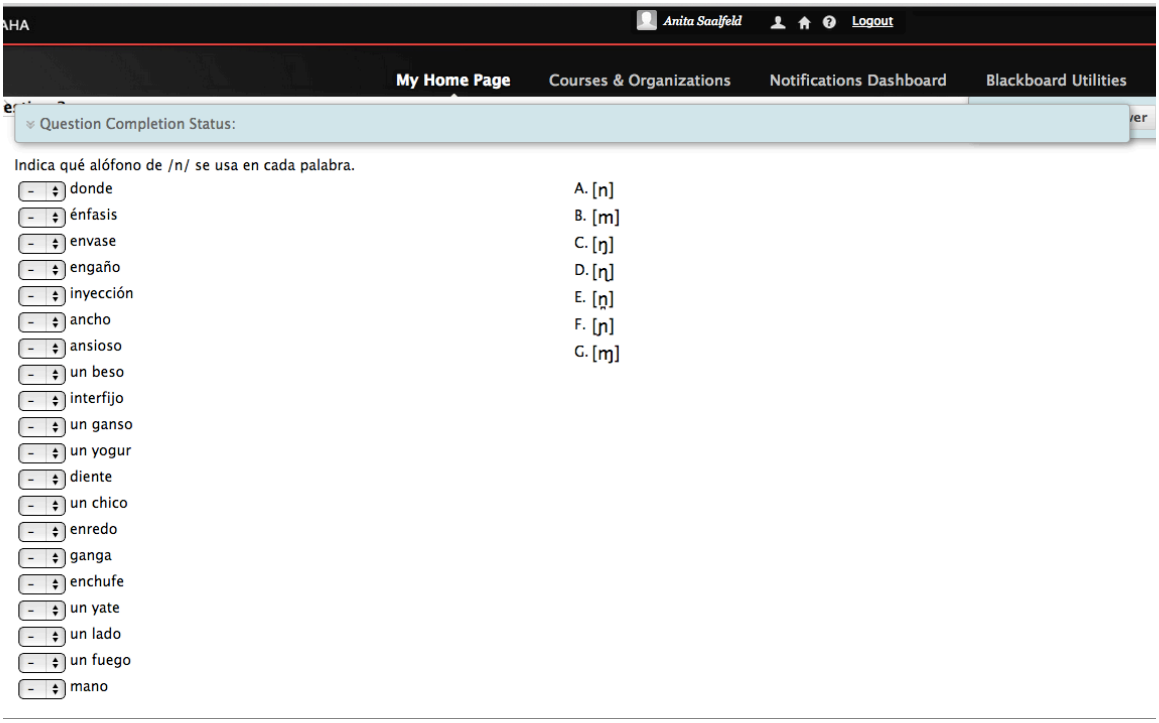


Figure 3. Matching question for allophones of /n/. Using a matching question rather than a multiple-choice or true-false question helps to make learners more aware of areas in which they need help.

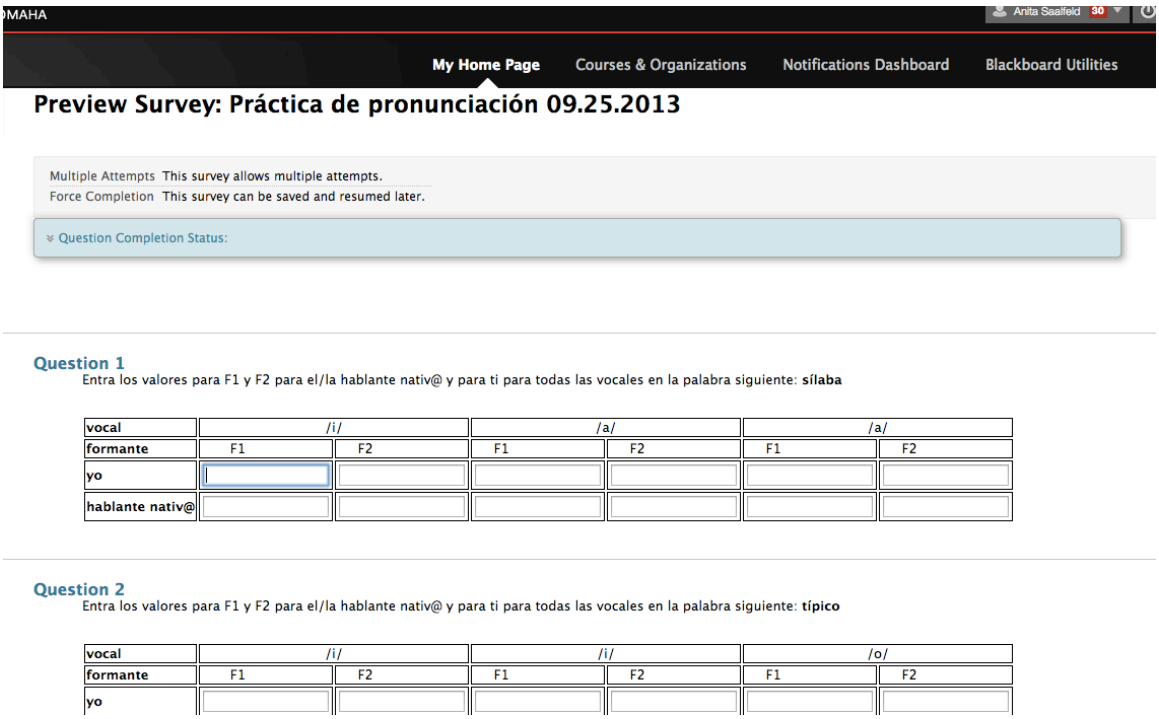


Figure 4. Pronunciation activity. This exercise required students to enter F1 and F2 values for both their own and native speaker productions of given words.

Meyers, C. (2014). Intelligible accented speakers as pronunciation models. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 172-76). Ames, IA: Iowa State University.

TEACHING TIP: INTELLIGIBLE ACCENTED SPEAKERS AS PRONUNCIATION MODELS

Colleen M. Meyers, University of Minnesota

Recent research has highlighted the benefits of choosing intelligible accented speakers as pronunciation models rather than native speakers of English. (Murphy, 2014). These models may be easier to imitate due to a variety of factors, such as slower speech rate, similarity in physical appearance, and models perceived as being attainable—not "perfect." Such models are particularly helpful for improvement in pausing, use of prominence, and use of non-verbal communication to portray confidence in speaking a second language. Finally, the skills learned by doing this project can be applied to more communicative situations in which the speakers will need to apply these pronunciation features, such as teaching, training, or other public speaking.

Goals of Teaching Tip

1. To help students identify pronunciation features in their own speech which prevent them from being effective communicators in speaking English.
2. To demonstrate the importance of choosing appropriate intelligible non-native speakers of English as pronunciation models.
3. To show the process of "mirroring" such a speaker, non-verbal language included, so that students can practice and apply those features to their own communicative context.

This teaching tip demonstrates how intelligible accented speakers of English can serve as successful models for L2 learners. The teaching tip moves students from awareness of their own pronunciation challenges, through identifying appropriate models, through the process of mirroring such a model to applying the skills they have learned to their own communicative setting. Tips for ensuring a successful project are also provided.

This teaching tip is an adaptation of projects done previously using native speakers as models. (Goodwin, 2004).

The particular example illustrated here is one in which the student needed to work on pausing for thought groups, prominence, and appropriate non-verbal communication. Thought groups correspond to grammatical units of speech and are marked by pausing and pitch changes. Prominence (or focus) in English is marked by pitch and length changes.

Prominence (also referred to as primary stress, nuclear stress, or focus) is a key feature in how well English speakers make sense of speech. Hahn (2004) looked at the use of prominence by international teaching assistants and she found that English-speaking listeners understood lectures better when prominence was correctly placed. Furthermore, the speakers who used correct prominence were evaluated more positively.

The speaker used to illustrate the process is a native speaker of Mandarin Chinese, preparing to work as a TA (teaching assistant) at the University of Minnesota. Several clips (Mary, Micro-teaching #1; Mary, Mirroring Project, Cold Version; Mary, Mirroring Project, Final Version) have been included to document the process (and progress) of moving from the beginning of the project to the end of it.

Step 1—Identify major pronunciation challenge

The first stage of this process is viewing the speaker communicating in an "authentic" environment. In this case, "Mary" (not her real name) was introducing herself and her class to her students on the first day of her MT (micro-teaching). As we can see, "Mary" has quite good control of her segmentals, but her suprasegmentals, in particular her use of thought groups, prominence and intonation are sorely lacking. Her speech sounds fairly monotonous, and without turning up the volume on the recording, hard to hear at times. In addition, if one considers the teaching context, "Mary" appears lacking in confidence as a teacher and rapport with her audience based on her lack of effective non-verbal communication, such as use of gestures, use of facial expression, and use of body movement. Rapport (Gorsuch, 2003) has been found to be one of the biggest and best predictors of success as a TA in the US.)

Step 2—Choose appropriate model.

The second step is to find an appropriate model, taking into consideration the challenge(s) identified in Step 1. In this case, "Mary" herself suggested Yang Lan, a non-native speaker of English and native Mandarin speaker who is known as the "Oprah of China."

Viewing Yang Lan, one can easily see why "Mary" chose her as the model. Yang Lan is expressive, charismatic, and confident. She speaks in clear thought groups with very good prominence. In addition, she engages her audience through her use of non-verbal communication, including expressive facial expressions, body movement, and hand gestures. Finally, her speech is told as a "story," making her topic even more engaging.

Step 3—Mark transcript for pronunciation feature(s) and non-verbal communication.

The next step is to transcribe the speech (if not already provided). A length of about 1 – 1.5 minutes (roughly speaking, 5 – 7 sentences or thought groups) is enough but not overwhelming. Less than this can be a waste of time; more can be too much and end up frustrating the students.

Many students choose speakers from Ted.com. There are several reasons, such as: 1) The speakers are well-known in their field, so they are highly articulate and express their ideas succinctly in English; 2) A transcript for each speech is provided—not only in English, but often in an array of other languages; and 3) One can click on the part of the speech under study to hear a particular sentence (or thought group) repeated over and over again. This last point makes the Ted.com recordings particularly well-suited to this type of project.

Many features of English pronunciation correspond to types of body language (Acton, 1984). Speakers will enunciate stressed syllables more clearly than unstressed ones. They will use hand or head movements to highlight prominent words, they will insert a short pause or "breath," when they want to move from one thought group to the next, and they often use eyebrow movements or facial expressions to underscore intonation patterns.

In this particular example, Yang Lan (Lan, 2011) uses several types of body language to highlight the words which she makes prominent, e.g., she raises her eyebrows when she says "final" and she stands tall and bends over, literally "mirroring" the intonation pattern of the phrase, "China's got talent show" to highlight it's in China (not America). Since the transcript was basically provided by Ted.com, "Mary's" job was to divide it into thought groups, locate the focus word (prominence) and then draw in by hand the body movement which corresponded to each focus word.

Step 4—"Mirror" the original recording one thought group at a time.

At this step, students need to familiarize themselves with both the spoken language AND the non-verbal language. Sitting at the computer, they play each thought group, pause and then mirror the phrase to the computer. They can do this in two stages: 1) focusing only on spoken language; and 2) adding the non-verbals to the spoken language.

Step 5—Practice using read, look up, and say technique.

Students already have their transcript from Step 3. At this point, they type the script in large font, one thought group per line. Then, they make the focus word bold. Finally, they draw in the non-verbal communication by hand.

Once this is done, they work in pairs. Each person goes one at a time. This first person will look down and read the first thought group, then look up and say the thought group to his/her partner. It's crucial at this step to actually make eye contact. This serves two purposes: 1) It forces students to incorporate a long enough pause to separate thought groups; and 2) It reminds students of how important eye contact is in speaking English to communicate sincerity and not just to robotically "parrot" words.

Step 6—Record "cold" version.

At this point, students should make a video-recording. They can either record during class or do so at home. The recording should be viewed, and students should be asked to identify places where they are doing well and areas which can be improved. They should document this in some way for the instructor. After that, the instructor can make additional comments. (At the University of Minnesota, we make use of Video Ant, an online program that allows users to insert markers at specific points on the recording and leave comments at each point marked.)

At this point, the recording may be good "technically," but it may lack emotion and/or the tone of the original recording. These aspects will be dealt with in the final step.

Step 7—Record "final" version with emotion.

Finally, a week or so later, students can make their final recording based on their self-assessment and feedback from their instructor. For this "final" recording, they should either memorize the script or write it in large font and put the paper somewhere they can see it easily. They can perform the script as many times as they wish, focusing on the "big picture" and not worrying if they make a simple mistake. Other students may be present and can serve as "directors" or audience members. Students often really enjoy this task and can serve as enormous help to each other.

To help them get into the "mood" of the original recording, students should talk about what the person is feeling in the original. Is the speaker enthusiastic? Sad? Angry? What is the speaker trying to accomplish? Teach? Inspire? Entertain? All of this makes the final product much richer and more beneficial for students.

Once the "final" version is recorded, it can be saved for the speaker and/or other classmates to view and learn from.

Tips for Success

In addition to the benefits of using Ted.com speakers, there are a few more tips worth mentioning.

First of all, all choices should be vetted by the instructor. One student in our class chose a speaker whose rhythm and intonation patterns made her unintelligible. This was actually beneficial because we could show the student why her choice was not a good one.

Typing the transcript in large letters, one thought group per line, is an effective and efficient way to help students get into the script. If they have a small hand-written script, they may end up just reading without emotion.

Getting the emotion right can be daunting in a second language. One technique which I've used with some success is to ask students to think of a time when they experienced this emotion, e.g., one of my students was trying to convince some people to vote his way. Only when I told him to imagine he wanted to convince US immigration to grant him permission to return to China to visit his ailing mother did he really "get" it.

Finally, viewing the recording with the sound off is a great way for students to observe the interaction between non-verbal communication and stress/prominence in English. It also helps facilitate the identification of emotion and tone. Students are often amazed at how much they can tell about what is being communicated simply by observing body language.

ABOUT THE AUTHOR

Colleen M. Meyers is a co-author of two of the most widely-used textbooks in the field of ITA education: Communicate: Strategies for International Teaching Assistants (1992, re-issued in 2007) and English Communication for International Teaching Assistants: Second Edition (2012). Colleen has presented both nationally and internationally at the Teachers of English to Speakers of Second Languages conferences, as well as at NAFSA, and the Pronunciation in Second Language Learning and Technology Conference. In 2007, Colleen was awarded a Senior Fulbright scholarship to do teacher-training (primarily pronunciation) at the Middle East Technical University in Ankara, Turkey. Colleen's primary interests include how non-verbal communication serves to enhance English prosody in academic communication.

REFERENCES

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly, 18*(1), 71 – 85.

Goodwin, J. (2004). The power of context in teaching pronunciation. In J. Frodesen & C.Holten, Eds. *The power of context in language teaching and learning* (pp. 225-236). Boston, MA: Thomson/Heinle,

Gorsuch, G. J. (2003). The educational cultures of international teaching assistants and US universities. *TESL-EJ*, *7*(3), 1-17.

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

Lan, Y. (2011). Yang Lan: The generation that's remaking China. Ted Talks. Retrieved from http://www.ted.com/talks/lang/en/yang_lan.html

Mary, Micro-teaching #1: http://mediamill.cla.umn.edu/mediamill/embed.php?subclip_id=3447&live=true&chrome=true

Mary, Mirroring Project, Cold Version: http://mediamill.cla.umn.edu/mediamill/display/156112

Mary, Mirroring Project, Final Version: http://mediamill.cla.umn.edu/mediamill/display/157042

Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL. *System*, *42*, 258-269.

Zhang, S. (2014). Using tongue twisters to supplement beginning level CFL students' pronunciation and tone practice. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 177-182). Ames, IA: Iowa State University.

TEACHING TIP
USING TONGUE TWISTERS TO SUPPLEMENT BEGINNING LEVEL CFL STUDENTS' PRONUNCIATION AND TONE PRACTICE

Shenglan Zhang, Iowa State University

Chinese is a tonal language and some sounds in Chinese such as <tɕ>, < tɕ^{h}>, <ɕ>, <ts^{h}>, and <ɥ>, do not exist in English. To learn tones and pronunciation well is a very challenging task for adult CFL learners whose native language is English. There are activities for students to train and improve their sensitivity on similar sounds and tones. For example, they can listen to and repeat the recordings of two-syllable words to detect the correct pronunciation or tones after hearing sets of words read out, and to speak in sentences and paragraphs in different contexts such as doing dialogues and performing with a well-written skit. These methods have been found useful, but they sometimes are a little boring. The teaching tip introduced below is an approach that could keep students more motivated to strengthen the accuracy of their pronunciation and tones: using tongue twisters to supplement beginning CFL learners' pronunciation and tone practice. Tongue twisters are phrases or sentences that were constructed to put similar but distinct phonemes and tones together to exercise the jaws, the tongue, and the muscles around the mouth. By purposefully putting the similar but distinct phonemes and different tones together, tongue twisters sometimes achieve comic effects. Because of these features, tongue twisters are challenging and engaging as well. This teaching tip demonstrates how tongue twisters can be used to exercise adult CFL learners' jaw, tongue and muscles around the mouth while enforcing their acquisition of the language structure and vocabulary and motivating students to engage in practicing tones and pronunciation. Although the tip talks about using tongue twisters in improving CFL learners' pronunciation and skills, I also give suggestions for ways that the principles behind the tip may apply to other languages.

Goals of Teaching Tip

1. To demonstrate how to incorporate tongue twisters in the curriculum.
2. To demonstrate how to select and sequence introduction of tongue twisters in a pedagogically appropriate way to improve pronunciation and tones maximally while enhancing learners' understanding of the structure of Chinese and expanding their vocabulary.
3. To demonstrate how to introduce tongue twisters so that students are actively involved.
4. To show how to involve students in creating tongue twisters.

Incorporating Tongue Twisters in the Beginning Level Chinese Curriculum

Tongue twisters are fun. However, if they are used very frequently, students will lose interest in them. In my curriculum for the beginning CFL learners, I introduce eight tongue twisters throughout the first semester. The first week of class focuses on giving learners a big picture of Chinese pronunciation (*pinyin*) and tones. Two tongue twisters are introduced in this week. The

two tongue twisters are very useful in helping students with tone learning and sound distinction. During the last 14 weeks, ten book chapters are learned and each chapter has different themes and introduces different grammar points (see *Integrated Chinese* Level 1 Part I for details.) After every two to three chapters, one or two tongue twisters are introduced based on either the theme of learning or the grammar points of the chapters. With only a few tongue twisters, students can practice each very thoroughly and keep their interest and enthusiasm.

Selecting and Sequencing Tongue Twister Introduction

Tongue twisters should not be selected at random. There are many well-known tongue twisters in Chinese. The selection needs to be based on the progress the students make in their learning so that the use of tongue twisters could not only help students with their pronunciation and tones, but also enhance students' learning of vocabulary and structure. I selected and modified some well known tongue twisters and also used made-up tongue twisters (by myself or my colleagues) to make them fit in what the students are learning at that time. Here I will explain why some tongue twisters are used at a certain time.

The first two tongue twisters introduced in the first week are[14]:

1. 妈妈骑马，马慢妈妈骂马； 妞妞骑牛，牛拗妞妞扭牛。
   Māma qí mǎ, mǎ màn māma mà mǎ; Niūniu qí niú, niú nìu niūniu nǐu níu.
   (Mom rides on a horse. She curses the horse because it's slow.
   Niuniu rides on a cow. She pinches the cow because it's stubborn.)

2. 四是四，十是十；十四是十四，四十是四十。
   sì shì sì, shí shì shí; Shísì shì shísì, sìshí shì sìshí
   (4 is 4, 10 is 10; 14 is 14 and 40 is 40.)

Tongue twister #1 is all about tones. When students are starting to learn Chinese without any background in the language, it is very crucial to make them aware of the importance of tones. Many characters are pronounced exactly the same, but they have different meanings with different tones. Tongue twister #1 conveys this message to the students. 妈妈(māma, mom), 马(mǎ, horse), and 骂(mà, to curse) have the same pronunciation, but they mean different things with first tone (mom), third tone (horse), and fourth tone (to curse). 妞妞(niūniu, a girl's name), 牛(niú, cow/ox), 拗(nìu, stubborn), and 扭(nǐu, to pinch) all have /niu/ sound, but the tones distinguish them as four different words.

Tongue twister #2 is introduced at this time for three reasons. First, the students learned numbers during the first week of the semester. A tongue twister with numbers in it can enhance students' memorization of numbers. In addition, the tongue twister containing both the number 14 and the number 40 could help students review how the teen numbers and how numbers counting by tens are formed. Second, /si/ and /shi/ are two similar but distinct phonemes that are very hard for some people in some regions of China to distinguish and sometimes create miscommunication in reality. This tongue twister can help students become aware of the diversity of Chinese language and culture. Third, the importance of tones is emphasized again. 十(shí, ten) and是(shì, to be)

[14] Most of these tongue twisters can be found in huayuworld.com website. See references.

bear the same pronunciation but have two different tones. And the second tone (十shí) and fourth tone (是shì) are also the tones that a lot of students are struggling with. This tongue twister gives students another chance to focus on the 2nd and the 4th tones.

Two new tongue twisters (#3 and #4) are introduced after the students learned measure words and time expressions in the first three chapters. In addition to practicing the similar but distinct phonemes /si/, /shi/, /zhi/, /zhe/, and /qi/, students are reminded of using measure words when numbers or demonstrate pronouns are used with nouns. For example, the measure word 只(zhī) must be used in front of 狮子(shīzi, lions) when a number (四十四, sìshísì, 44) is used with the noun in #3. In #4, the measure word个 (ge, the most commonly used measure word) must be used after the demonstrative pronoun 这(zhè, this) and in front of 星期四 (xīngqīsì,Thursday). Tongue twister #4 comes at the right time when the students just learned that the time expression must be put in front of the action in Chinese. The time phrase (星期四, Thursday) is in front of the verb phrase (有事, to have commitment). Word order is so important in Chinese that any opportunity to remind students of the word order is valuable, especially right after they learned a new word order.

3. 四十四只石狮子死了。
   sìshísì zhī shíshīzi sǐ le.
   (44 stone lions died.)

4. 石老师这个星期四有事。
   shílǎoshī zhège xīngqī sì yǒu shì.
   (Teacher Shi has commitment this Thursday.)

After learning about eating and drinking in chapter 5 and chapter 6, I introduce to students a classic tongue twister, tongue twister #5. This one could help expand students' vocabulary related to juice/food (such as 葡萄grapes, 葡萄皮 grape skin) while enhancing students' knowledge about tone sandhi when using 不(negation word), that is, when 不(bù) is followed by 吐(tù) as in the first clause, it changes to second tone (bú). However, when it is followed by the first (the word 吃chī in the second clause), the second or the third tone, it should be the fourth tone as /bù/.

5. 吃葡萄不吐葡萄皮，不吃葡萄倒吐葡萄皮。
   Chī pútao bú tù pútao pí, bù chī pútao dào tù pútao pí.
   ([He] eats grapes but does not spit the skins out. [He] does not eat grapes, but spits out grape skins.)

When practicing tongue twisters #6 and #7 in chapter 9, which introduces colors, students not only can practice the sounds /h/ and /f/ that are very hard to pronounce when being put together with /eng/, /ang/, /ei/, and /hui/, but also can review and expand the color words, such as 粉(pink), 粉红(pinkish red), 灰(gray), and 黑灰(dark gray).

6. 红凤凰，粉凤凰，粉红凤凰飞。
   Hóng fènghuáng, fěn fènghuáng, fěnhóng fènghuáng fēi.

(Red phoenix, pink phoenix, pinkish red phoenix fly.)

7. 黑蝴蝶飞，灰蝴蝶飞，黑灰蝴蝶飞。
   Hēi húdíe fēi, hūi húdíe fēi, hēihūi húdíe fēi.
   (The black butterfly flies, the gray butterfly flies, and the dark gray butterfly flies.)

The last chapter in the book is about transportation. One of the very important grammar point is that means of transportation should be put in front of the verb phrase, as the time expression (see tongue twister #4). Tongue twister #8 is introduced with the grammar point in mind. 坐飞机(zùo fēijī, to take airplane) is the means of transportation and it is put in front of the verbal phrases 去斐济(qù fēijì, to go to Fiji) and 吃肥鸡 (chī féijī, to eat fat chicken).

Meanwhile the importance of tones is emphasized for the last time in the semester. 飞机(fēijī, airplane)，斐济(fēijì, Fiji)，and 肥鸡(féijī, fat chicken) have the same pronunciation but have totally different meanings with different tones. The meaning of this tongue twister is hilarious and it could make a happy ending when it is used as the last tongue twister for the semester.

8. 坐飞机去斐济吃肥鸡。(created by Dr. Tonglu Li)
   zùo fēijī qù fēijì chī féijī.
   (Go to Fiji by airplane to eat fat chicken.)

Introducing Tongue Twisters

When introducing tongue twisters to students, the principle I follow is to get involved as much as possible. I usually follow the following three steps. First, I ask the students how to write certain words that appear in the tongue twister. If there are words that they do not know, I will write the characters on the blackboard. I say how they are pronounced and then ask the students to write down the *pinyin* and tones for the characters. The students will pronounce the individual word for a few times making sure they know each word very well.

Once all the words are introduced, I ask the students to make them into meaningful sentence(s). Sometimes I put the words into sentence(s) and ask the students to tell me the meaning. After making sure that the students understand the structure and the meaning of the tongue twister, I ask the students to read it slowly. Funny images illustrating the meaning of the tongue twisters (when PowerPoint is used) are often used to make the activity more interesting.
In class, the students will be given some time to practice. Students can volunteer to record their tongue twister practice (using the instructor provided audio-recorder) and ask the instructor to replay and correct it with the whole class. Outside class, the students will be required or encouraged to practice and record his/her saying of the tongue twister and publish it at Voice Board at Blackboard Learn.

Encouraging Students to Produce Tongue Twisters

Students have occasionally been encouraged to create their own tongue twisters in groups based on the vocabulary and grammar they learned. Each group presents their tongue twister in class

and gets feedback from both the instructor and their peers. Then each group practices and post their recorded tongue twister at Voice Board at Blackboard Learn. Some of them created very interesting ones. See the examples, tongue twisters #9, #10, and #11.

9. 他是老师不是律师。
   Tā shì lǎoshī búshì lüshī.
   (He is a teacher, not a lawyer.)

10. 你哪个哥哥喜欢唱歌?
    Nǐ nǎge gēge xǐhuān chàngge?
    (Which older brother of yours likes singing songs?)
11. 我想坐出租出去吃饭。
    Wǒ xiǎng zuò chūzū chūqü chīfàn.
    (I want to take a taxi to go eating.)

I suggest that a variety of awards such as the best pronunciation award, the best creation award, the most fun tongue twister award, etc., be given to motivate students to practice the given tongue twisters and create their own.

Final Thoughts about Using Tongue Twisters in Language Teaching

There are two principles for using tongue twisters in language teaching. First, quality is more important than quantity. The best tongue twisters are those that could not only help students practice pronunciation and tones in a fun way but also help them reinforce the grammar, vocabulary, and culture they have just learned. Second, find interesting ways to get students involved, such as using fun images to illustrate the meaning, asking students to put a tongue twister together, and/or giving rewards to encourage creation.

Tongue twisters can be used in teaching other languages. Teaching pronunciation in foreign/second language education can be exhausting. The most commonly used repetition drills could make the lesson monotonous. Using tongue twisters, however, can make the pronunciation teaching and learning fun. Students are less afraid of making mistakes when everyone including the teacher might make mistakes. Besides, tongue twisters can be integrated in teaching other language skills (vocabulary, structure) and culture while teaching pronunciation. For example, the following popular English tongue twister can be used in multiple ways in teaching English.

Peter Piper picked a peck of pickled peppers.
Did Peter Piper pick a peck of pickled peppers?
If Peter Piper Picked a peck of pickled peppers,
Where's the peck of pickled peppers Peter Piper picked?
(Fun-with-words, 2014)

There are a few teaching points when this tongue twister is used. The pronunciation of /p/ followed by different vowels can be exercised. In addition, past tense, attributive clause and the use of the article “the” and “a” can be practiced. Lastly, the tradition of American pickled vegetables can be introduced through this tongue twister.

ABOUT THE AUTHOR

Shenglan Zhang, Adjunct Assistant Professor in the Department of World Languages and Cultures and Lecturer in the School of Education at Iowa State University, received her Ph.D inEducational Psychology and Educational Technology from Michigan State University and her M.A. in East Asian Languages and Cultures from the University of Southern California. Her research interests fall within three different but related areas: integrating technology in teaching Chinese language, blended and online teaching, and reading on the Internet. Currently she is working on two projects, one exploring Chinese-as-a-foreign-language (CFL) learners' use of Chinese with native speakers within a computer-mediated format, and the other examining the use of digital story-telling to support CFL learners' speaking. She can be contacted at shenglan@iastate.edu.

REFERENCES

Fun-with-words website, retrieved on July 21, 2014 at http://www.fun-with-words.com/tong_example.html

Huayuworld website, retrieved on June 20, 2013 at http://blog.huayuworld.org/feeling7866/19158/2010/10/03/74504

Roccamo, A. (2014). Effective pronunciation instruction in basic language classrooms: A modular approach. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 183-189). Ames, IA: Iowa State University.

TEACHING TIP
EFFECTIVE PRONUNCIATION INSTRUCTION IN BASIC LANGUAGE CLASSROOMS: A MODULAR APPROACH

Ashley Roccamo, The University of Southern California

Pronunciation skills in a second language (L2) are a significant component of communicative competence, yet are notoriously difficult to acquire. Most L2 learners seem unable to identify and improve their pronunciation on their own (Derwing & Munro, 2013; Derwing & Rossiter, 2002; Dlaska & Krekeler, 2008); therefore it seems necessary to provide them with instruction in pronunciation. Instruction programs have been shown to be effective in improving pronunciation skills, by reducing accent and increasing comprehensibility (e.g., Derwing, Munro & Wiebe, 1998; Elliott, 1995; Hardison, 2004; Lord, 2008; Saito & Lyster, 2012). Yet it remains uncommon to encounter a pronunciation training study implemented in beginner or intermediate four-skills classrooms, despite the fact that researchers are beginning to promote the idea of beginning pronunciation training as early as possible (Counselman, 2010; Derwing & Munro, 2013; Eskenazi, 1999; Munro, 2013). This is not for lack of concern, however; many foreign language instructors recognize the importance of pronunciation skills, but do not teach them because they are worried about time constraints or are unsure of how to get started.

The following article provides information to counteract both of those worries, providing suggestions for a module-style pronunciation instruction unit that uses warm-up activities to teach pronunciation in just ten minutes per class period. This modularized unit has been tested with L2 learners of German and was successful at providing them with the tools to improve their comprehensibility and reduce their accentedness to a greater extent than learners who did not receive training. The first section of this article explains the motivations behind using a modular design. The basic structure of each module is outlined in the second section, and information about classroom instruction and sample activities can be found in the third. It is hoped that interested instructors can use the information found here as a springboard to introduce similar modules into their own basic language classrooms.

Incorporating Pronunciation Instruction into Beginner-Level Classrooms

Three things are of utmost importance when incorporating pronunciation instruction into basic language courses: the issues of time, adaptability and teacher confidence. So much needs to be learned in beginner and intermediate four-skills language classrooms that instructors are often pressed for time, which is why a modular design is useful. Students receive two weeks of intensive instruction on one aspect of pronunciation without taking too much class time away from grammar, vocabulary and cultural lessons. Within each module, pronunciation activities are designed to be used as a warm-up activity for the first ten minutes of class each day, leaving the rest of class time free. Ideally, pronunciation activities in the module combine the practice of many language skills with pronunciation practice, including speaking, listening, reading, writing,

grammar and vocabulary. The activities used in the pronunciation modules can also be custom-tailored to correspond to the course themes currently being covered. Instructors can easily combine a review of what was just learned with a pronunciation warm-up before continuing with a new topic.

Another important feature of the modular pronunciation unit is its adaptability. Each module is self-contained, which allows for instructors who are particularly pressed for time to implement the modules in an a-la-carte manner. Instructors can choose one or two modules to focus heavily on the needs of their particular courses and students. Once the basic format of the module is learned, it can be adapted to whichever languages, course goals, or first language (L1) groups an instructor encounters.

The pronunciation training unit described here also addresses the issue of teacher confidence. Many foreign language instructors who were not trained in linguistics or phonology may feel that they simply do not know enough to teach pronunciation. The activities involved in a unit of this type are designed to avoid technical phonetic instruction, focusing instead on familiar classroom activities that instructors already have practice implementing. Instructors can thus help their students improve their pronunciation without getting caught up in terminology and work within a realm that feels comfortable to them.

Module Design

When designing modules for use in the basic-level language classroom, it is vital to provide L2 learners with two skills: 1) the ability to produce what is currently being trained in a more accurate manner, and 2) the ability to notice and attend to their own pronunciation and that of others. Through this type of instruction, students can improve their pronunciation by noticing the differences between what they produce and what native and advanced speakers do, and by learning how to actually produce the new or similar sounds themselves.

A number of factors are involved in setting up modules for a particular course. Each module lasts for two weeks and focuses on one aspect of pronunciation that needs attention. There are no limits to what can be trained; instructors should choose features based both on students' needs and L2 problem areas typical for learners with a certain L1. Research findings recommend that at least one prosodic aspect be chosen for focus, such as placement of stress or intonation, as errors in prosody can be more detrimental to comprehensibility than errors in individual sounds (Hirschfeld, 1994). Once the topic for each module has been chosen, instructors can design pronunciation activities that can be used as a warm-up for the first ten minutes of class time each day.

Training in both perception and production should be provided, as well as activities that heighten attention and noticing of the feature in focus. The activities in each module are intended to follow the optimal progression for pronunciation training as outlined by Chun (2002). Students begin each training module with active listening exercises and end with production exercises. Listening and perception training consists of awareness training, which outlines what the feature is and how it is used in the L2, followed by discrimination and identification training, respectively. In discrimination training, students hear two words and decide whether these words

were the same or different. Identification training is when students hear a word and must use their knowledge to decide exactly which word it was that they heard. In all perceptual training activities, students should be provided with authentic speech samples from many different speakers that vary in length from individual words to longer dialogues and texts.

After thorough perceptual training, students should then begin practicing how to produce the individual pronunciation features. This stage makes selective use of traditional pronunciation training methods, such as drills and mimicking, to help students realize what their mouth is doing when they produce new sounds. Afterwards, students will progress to more active controlled speech exercises that practice the targeted features in a variety of different contexts. During this stage, students will work from more simple utterances, such as those with just one or two words, through sentence- and paragraph-level speech to spontaneous speech contexts that are modeled after natural language situations in which their L2 would be used.

As designed, each module begins with an activity to focus attention and heighten awareness of the aspect being trained and ends with a free speech activity. A sample of the progression of each of the activities is outlined below.[15]

Day 1: Awareness Training
Days 2-3: Discrimination Training
Day 4: Identification Training
Day 5: Production Training—Individual Sounds/Individual Words
Day 6: Pronunciation in Individual Words/Sentences
Day 7: Pronunciation in Sentences and Paragraphs
Day 8: Free speech practice

The progression of the module is important because it addresses two L2 problems that influence pronunciation skills: perceptual skills and production of the targeted feature. Chun (2002) argues that it is important to first aid the development of L2 perception and listening skills before beginning any practice with production. Production activities move from more simple to more complex in order to promote stabilization of the speech patterns that are being learned. Ending with spontaneous speech contexts provides room for instructors to incorporate plenty of communicative language practice that demonstrates the importance of the pronunciation feature in question for communicating meaning.

A key component of these pronunciation modules is the concept of noticing. Throughout all of the modules and activities, students' awareness both of the targeted features and differences in pronunciation between the L1 and L2 should be raised as much as possible. Form-Focused Instruction can make use of enhanced input in communicative contexts to draw students' attention to the relevant features and how they affect comprehensibility, intelligibility, and meaning. Students should focus on the pronunciation of their peers as well; research has shown that by attending to another students' pronunciation and offering feedback, students can improve

[15] The pronunciation unit was originally designed to take place in a classroom that meets four days a week, but could easily be shortened or extended. In addition, less time could be spent on perception training if students seem to have no problems with it, in order to spend more time on production training.

their own (Counselman, 2010; Lord, 2008). Explicit examples of activities making use of peer feedback and partner work within the frame of the training progression are outlined below.

Classroom Instruction and Example Activities

In the classroom, pronunciation instruction should be completed as a warm-up at the beginning of the class period. The sample activities presented in this section are meant to combine pronunciation practice with various language skills. Each of the following activities can be adapted to train a variety of segments and prosodic patterns.

The first three activities presented are listening and perception activities meant to be used in the first week of each module. Activities four through seven provide examples of production training in which students practice producing target features in many different communicative situations and discourse levels. Many of the production activities require partner work so that students can test their listening and speaking abilities within the same activity. As mentioned previously, critical to all partner work is the aspect of feedback.

Awareness Training
Sample Activity: Listening to a Text

For this activity, students listen to a text (poem, short story, song) and read along. Target sounds or prosodic features are made prominent in the text in order to focus students' attention. Attention can be focused on orthographic-phonetic connections, the possible sounds that can be used in different contexts, or differences between L1 and L2 pronunciation of similar sounds. After the listening phase, the instructor can ask students about the targeted features and their representation in the text. The primary goal of this activity is to raise students' awareness about the pronunciation of a certain feature in the L2, its characteristics, and how it is marked orthographically. Authentic texts are especially useful here, as they can be worked into the rest of the lesson after the pronunciation exercise is completed.

Discrimination Training
Sample Activity: Discrimination Game

Students playing a discrimination game will hear a pair of words pronounced out loud. Their task is to decide whether the pair consists of the same word said twice or two different words. Students can mark their answer in a number of different ways: by circling "same words" or "two different words", by checking a box, or by responding with a clicker or computer. A very useful way for instructors to see who can hear the differences and who cannot is to have students hold up cards with one color for "same" and another for "different". Instructors can make the back of the cards white so that students will not feel shy or wait to see the others' answers.

Instructors should choose words for this game that are identical except for the sounds that are being contrasted, such as the example *thin* and *tin*. This is a particularly versatile game, as it can be adapted to stress patterns and individual sounds in any language that has phonological contrast. Words can be contrasted both within and across languages. For instance, one could contrast short and long vowels within English words (e.g., *bit* and *beat*), or one could contrast the

diphthongized American English pronunciation of /e/ with its pronunciation in Spanish or German, which both make use of the monophthong (e.g., English *zone* and German *Sohn*). When creating a discrimination game, an instructor must make sure to include not only pairs that are different, but also pairs that are the same. Goals of a discrimination game are raising awareness of the contrasts between two L2 sounds or prosodic features, or between L1 and L2 pronunciation of a similar feature. This type of activity can also be adapted to intonation patterns, such as statements and questions.

Identification Training
Sample Activity: Bingo

Identification and discrimination are similar tasks, because they can be adapted to train similar contrasts and require similar responses from students. Yet there is one major difference—in identification games, students hear different words pronounced and have to indicate *which* word or sound was heard, instead of just deciding if the words were the same or different. One creative type of identification game is Bingo. In Bingo, each student receives an individual Bingo board. They must listen carefully to words pronounced aloud and then put a marker on the correct word on their Bingo sheet. This is a great game to help students listen for contrasts between similar sounds, and is also a useful tool for language instructors, as recent vocabulary items can be tested at the same times as valuable L2 contrasts. Goals for identification games are to test contrasts between the students' two languages, to raise awareness of the differences in those contrasts, and to train students' perceptual and word identification abilities.

Producing Individual Words
Sample Activity: Telephone

Once students have practiced some critical words involving the target pronunciation feature, they can play "Telephone"[16]. Students split into two or three groups, depending on the size of the class, and form lines from the front of the classroom to the back. The student at the head of the line receives a slip of paper with a word on it that uses the targeted contrast. They must say this word to the next person in line, who must say it to the next person, and so on until the end. If the last person in line produces the same word that the person at the beginning of the line has on the paper, that team wins a prize. In this game, students must both listen carefully to the relevant contrasts and pay close attention to which sounds they are forming as they speak.

Pronunciation in Sentences
Sample Activity: Question and Answer Sessions

In this controlled production activity, each partner receives a list of questions and a list of answers. Their answers correspond to their partner's list of questions, not their own. Students listen to their partner ask the questions, and choose one of their answers in response. Both partners give feedback to the other after they have asked and answered all questions.

[16] This is called "Whisper Down the Lane" or "Operator" in some regions.

This activity is particularly useful for highlighting the connection between pronunciation and meaning, as students must listen closely to their partners' pronunciation before they can choose an appropriate response. Having students attend to their partners' pronunciation can be beneficial for them as well. This activity is particularly well suited to modules that focus on sentence accent and contrastive intonation patterns, but is useful to train individual sounds in questions and sentences as well.

Pronunciation in Paragraphs
Sample Activity: Poetry Reading

After practicing their pronunciation in sentences and questions, students can move on to larger texts, like poems. In this sample activity, students receive a poem and practice performing it while focusing on their pronunciation. The poem may be the same as was used for the first day's listening activities, and the target features can be made prominent in the text once again. This activity can be given a competitive slant, wherein students listen to and critique each other's performances. One could also have intermediate students write their own poems before they perform them for a partner or the class. Goals for this activity are to be able to focus on pronunciation and accurately produce the targeted feature even as the discourse level becomes more complex.

Free speech practice
Sample Activity: Describing a Picture

It is important to end a module with at least one free speech activity, in order to practice pronunciation of the target features in a complicated speech situation. Only one activity is described here, but instructors and students have great freedom when designing and completing free speech tasks. The connections between other lessons in the classroom and pronunciation are strongest with the free speech activities; culture can be incorporated by using authentic materials from the target culture, and recently learned vocabulary or grammar topics can be reviewed.

For this sample activity, each student receives a picture depicting something related to their chapter lessons and they take turns describing it in some way. Based on ability level, students may name the objects and actions they see, talk about their physical properties, continue the story beyond what is pictured, or make cultural comparisons. This activity can also easily be turned into a game where students describe something from the picture and their partner has to guess what it is. What is key is that students are practicing the target pronunciation features while they are actively using their L2 without reading aloud.

Conclusion

The module-style pronunciation instruction unit described above is unique in that it teaches pronunciation in just ten minutes per class period, and is a successful technique for teaching beginner- and intermediate-level students in basic language classrooms. The guidelines outlined here can provide instructors with a springboard to overcome their concerns about implementing pronunciation instruction in four-skills classrooms, including worries about comfort level or time restrictions. It is hoped that through this short tutorial, instructors who value pronunciation skills

in L2 learning can begin to recognize both the feasibility of providing pronunciation instruction in beginner- and intermediate-level basic language classrooms and can feel more comfortable getting started.

ABOUT THE AUTHOR

Ashley Roccamo received her Ph.D. in German Applied Linguistics with a Dual Title in Language Science from Penn State University in May 2014. She is currently the Associate Director of Articulation and Assessment at the University of Southern California Language Center. Her research interests lie in the effects of instruction on second language German pronunciation, particularly for learners in their first few semesters of language learning. Secondary interests include: the role of motivation in the acquisition of pronunciation skills and language contact between English and German. She has presented at multiple conferences, including at the Germanic Linguistics Annual Conference, New Sounds, and PSLLT, and has a publication forthcoming in *Die Unterrichtspraxis*.

Reed, M. (2014). The English syllable: Big news, bad news and its importance for intelligibility. In J. Levis & S. McCrocklin (Eds). *Proceedings of the 5th Pronunciation in Second Language Learning and Teaching Conference* (pp. 190-203). Ames, IA: Iowa State University.

THE ENGLISH SYLLABLE: BIG NEWS, BAD NEWS, AND ITS IMPORTANCE FOR INTELLIGIBILITY

Marnie Reed, Boston University

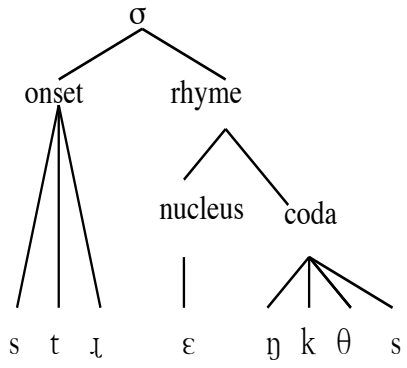
Addressing bi-directional intelligibility challenges: From the humble syllable to the pragmatic functions of English prosody

The negative impact of non-native prosody on speaker intelligibility has been documented both when segmental accuracy is intact (Derwing & Rossiter, 2002) and when it is compromised (Zielinski, 2008).

The challenge for L2 learners when interpreting native-speaker prosody has also been documented (Vandergrift & Goh, 2012) by learners who report, "understanding the words but not the message" (p. 22). This paper seeks to address instructor and learner metacognition regarding the role of the English syllable at the lexical, phrasal, sentential and discourse levels. It includes monosyllabic native-language/ target-language syllable-structure comparison. It addresses the spoken form category of Nation's (2001) requirements for knowing a word by adopting the Murphy and Kandil (2004) numeric notation system for lexical stress. It expands the scope of this system to the phrasal and sentence levels. Finally, it establishes a springboard for recognizing contrastive stress and intonation and their role in conveying speaker intent.

Bad news

Of 486 languages recently surveyed, English is reported in *The World Atlas of Language Structures* to be among the roughly 30% classified as having a complex syllable structure (Maddieson, 2013). The complexity of the English syllable, which allows onsets of up to three consonants phonemically and codas of up to four consonants phonetically in monosyllabic words – for example, *strengths*, diagramed below – poses challenges to learners from simple or moderate syllable structure languages.



Challenges that the complexity of the English syllable poses are met one of two possible ways: 1) final consonant deletion or consonant cluster reduction, eliminating some or all coda consonants, or 2) epenthesis, inserting vowels to restore a C-V syllable structure. These phonological processes seem to be used unconsciously as learners seek to conform to their L1 syllable structure. Both solutions adversely impact intelligibility, either singly or in combination with morpho-syntactic errors:

Thai speech sample:

	sounds like
You like white rice?	You lie why rye?
CV CVC CVC CVC	CV CV CV CV

Chinese[1] speech sample, spoken by a tailor to a customer seeking to get pants hemmed:

One in(ch) or two in(ches) ?
CVC VC VC CV VC

Japanese speech samples:

gift shop	gifuto shoppu[2]
CVCC CVC	CVCVCV CVCV
hot dog	hoto dogu
CVC CVC	CVCV CVCV

Teaching tips

Take advantage of target sounds/ clusters that exist in the L1. The problem is not always with the articulators; target sounds may exist in the L1 but not in all possible word positions or phonological environments. Examples: Thai has unvoiced alveolar fricative /s/ word initially (sà-wàt-dee), but not word finally; Spanish allows clusters with fricative /s/ + stops word-internally ("transcriber"), but not word-initially ("escribir"); unvoiced bilabial stop /p/ occurs in Arabic as an allophone [p] of /b/ before unvoiced consonants. Bypass orthographic mapping, word position, and other L1 constraints. Migrate permissible articulations to new word positions and phonological environments.

Start contrastive stress instruction early. Introduce marked, contrastive stress with monosyllabic words, for example in lessons on antonyms. Anticipate push-back; students may resist exaggerated prosodic contours of the target language. Assuming that speech production facilitates speech perception in an auditory feedback loop (Reed & Michaud, 2012), encourage students to practice producing contrastive stress in class in order to hear it when used by others outside of class.

I don't want a *small* one, I want a _____ one.
Don't turn the volume *up*, turn it _____.
your turn: _____.

Important news & more bad news

Moving on from monosyllabic words, lexical stress poses a second challenge, establishing the syllable as the relevant and meaningful unit of timing for English, the duration of the vowel in the stressed

[1] Chinese and Japanese allow CVn syllables, e.g., Japanese *San* for "Mr./ Mrs./ Ms."; Chinese [kwaŋ] for "light"

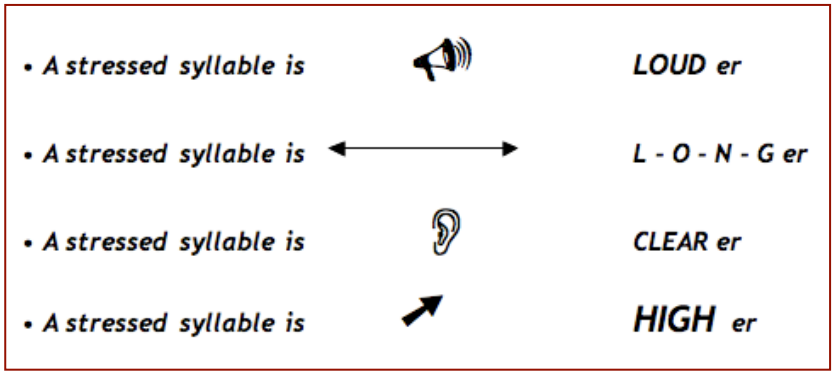
[2] Gift Shop at Golden Gate Park, San Francisco. Photo courtesy of J. Gilbert

syllable as essential for intelligibility, and the alternation of stressed and unstressed syllables as the foundation for English prosody. The important news is that stress alternates in most disyllabic and polysyllabic words[3] (see Cutler, 1986). The bad news is that lexical stress patterns are not easily predictable, as confirmed by the identification of 39 stress patterns in an analysis of 2979 polysyllabic words in Coxhead's (2000) Academic Word List (Murphy & Kandil, 2004). Lexical stress in di-syllabic words: In English, stress patterns alternate.

trochaic: **STRONG** weak	iambic: weak **STRONG**
differ	defer
ego	ago
awkward	occurred
person	percent

The phonetic realization of stressed syllables includes dynamic cues, melodic cues, and vowel quality changes (Sluijter & Van Heuven, 1996).

Perceptually, the salient characteristic of stressed syllables is prominence (Roach, 2009), the recognizable features of which include the following:



Challenges

In English, lexical stress patterns in most polysyllabic words are unpredictable.

Three 3-syllable words;
Three different stress

Wind Instrument	piccolo
Keyboard Instrument	piano
String Instrument:	violin

We know that word stress is important for spoken word recognition (Cutler, Dahan, & van Donselaar, 1997), as illustrated below:

[3] Exceptions: compounds (*sidewalk*) and polysyllabic words with full vowels (e.g., *halo*, *window*, *typhoon, hotel*)

Linguistics Conference Discussant: "I'm here to discuss the [wa ka .**byu** la riz]."
vocabulary

Workplace Training Session: "The # 1 priority is control of the [In .**wɛn** to ri]."
inventory

Call Center Customer Service Rep: "To activate your cell phone, push a row."
the arrow

We also know that for L2 learners of English, word stress is often reported to be difficult to acquire (Archibald, 1993, 1997; Guion, 2005; Guion, et al., 2004; Pater, 1997; Wayland et al., 2006). The degree of predictability in the L1 plays a role in cognitive attention to word stress; learners from fixed-stress languages never have to think about word stress.

1st Syllable: Czeck, Finnish, Icelandic, Hungarian
Penultimate Syllable: Quechua, Polish
Antepenult Syllable: Macedonian
Word-final Syllable 'prosodic stress': French

According to Peperkamp, Vendelin, & Dupoux (2010), if word stress is predictable, learners do not encode it in their lexical representations in their L1, nor, crucially, when learning an L2.

Teaching tips

In variable-stress languages, such as English, the number of possible stress patterns is unwieldy. As a manageable alternative to teaching a large number of rules, provide tools.

Stress Pattern Notation System*
Two numbers:
The first number indicates the number of syllables
The second number indicates where the primary stress falls.

3-1	3-2	3-3
piccolo	piano	violin

*adapted from Murphy, J., Kandil, M. (2004). Word-Level Stress Patterns in the Academic Word List. *System, 32,* 61-74.

Alternatively, you may wish to adopt a Stressed Syllable/Number of Syllables system, rendering the above: piccolo – 1/3 (1st syllable stressed in a 3-syllable word), and so on.
A former student, schooled in the Murphy /Kandil notation system, added colored squares as a visual aid.

Screen shot of classroom set-up courtesy of M. Noble, ESOL Instructor, Salem, MA

It is fine to let your students see you counting out the syllables on your fingers. Native Speakers have long forgotten how we acquired the stress patterns along with every multisyllabic word we know.
Once this system is introduced, resist responding to, "How to say ___?" by saying the word. Rather, provide the stress pattern and let the students produce the word. They will retain it longer – and they love this system!

Embed the stress-pattern notation system in a Vocabulary Checklist (Reed & Michaud, 2005, p. 154). Introduce it once; use it thereafter. It addresses the question, "What do your students need to know when they learn a new word?" in keeping with Nation's list for receptive and productive skills (2001, p. 27).

Learning New Vocabulary

When you want to know the meaning of a word, you need to ask the question grammatically:
Asking someone what a word means: What does _____ mean?
Telling someone what a word means: It means _____.
Did you know?
When you learn a new word, you need to learn its Stress Pattern as well as its meaning.

New Word Stress Pattern

For example: *economy* is a 4.2 word (4 syllables, stress on the 2nd syllable)
economics is a 4.3 word (4 syllables, stress on the 3rd syllable)

Checklist for Learning New Vocabulary

1. How do you spell _____?

How do you spell it?

2. What part of speech is it? **_____**
For Nouns: Count Noun? ☐ Non-Count Noun? ☐
Singular Count Noun: a/ an/ the
Plural Count Noun: add 's'
For Verbs: Transitive? ☐ Intransitive? ☐

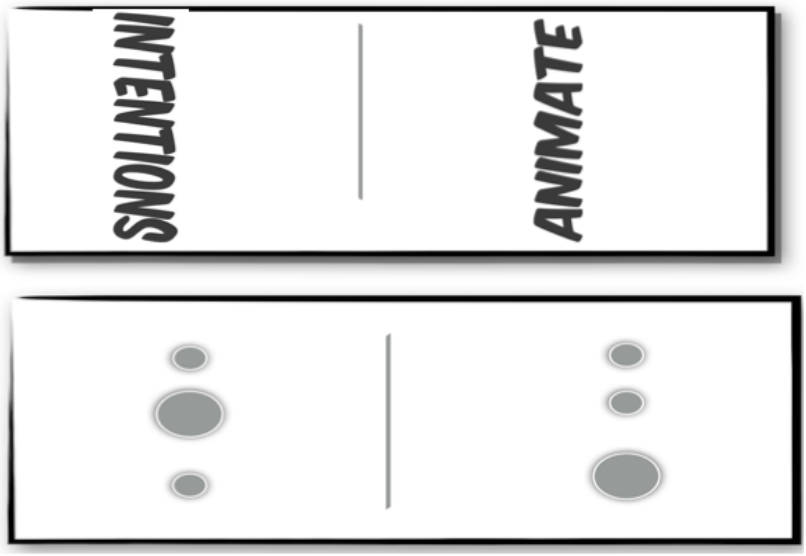
3. How do you pronounce it?
How many syllables are there in the word? _____
Which syllable gets the (primary) stress? _____
New Word / Stress Pattern: _____ / _____

4. How do you use it in a sentence?
_____

5. Alternate Forms:
_____

Teaching Tips Dominoes

Provide Plentiful Practice Opportunities. Once the lexical stress notation system for multisyllabic words has been introduced, provide opportunities for students to practice at both metacognitive and oral skills levels. Create and use Word Dominoes to focus learners' attention and promote practice with stress patterns in multisyllabic words. Include commonly mispronounced monosyllabic words (<speak> not [ɛspik]; <lunch> not [lʌntʃi], and regular monosyllabic nouns and verbs with inflectional endings (-es; -ed), which students typically mispronounce as two syllables: (e.g., <fixed> [fiksɛd]).



Word Dominoes
Inspired by Celce-Murcia, et al. (2010). *Teaching Pronunciation: A Course Book and Reference Guide.* Cambridge University Press.

Adapted from Hancock, M. (1995). *Pronunciation Games.* Cambridge University Press.

Big news

As noted by Cutler (1986), in English, "stress oppositions between verb and noun forms of the same stem {*decrease, conduct, import}* are common…" (p. 204). Exercises which highlight these oppositions, such as the one below adapted from Reed & Michaud (2005, p. 61), can be eye-opening for learners. This may be the first time learners notice and recognize a function of stress: Change the stress pattern; change the part of speech. At the metacognitive level, learners can make observations and generalizations about stress assignment: trochaic stress assignment to nouns; iambic to verbs. Teachers can optimize this 'ah-hah' moment and maintain a focus on normal and also contrastive stress.

		Stressed Syllable	Part of Speech
1.	a) What an insult!	1st	Noun
	b) Don't insult me!	2nd	Verb
2.	a) Round up the suspects.	1st	Noun
	b) I think he suspects you.	2nd	Verb
3.	a) Do I need a permit?	1st	Noun
	b) We don't permit that.	2nd	Verb

Teaching tips

When teaching derivational affixes, teach normal, predictable suffix stress patterns first.

trochees (**STRONG** weak)	+	-ation
Primary stress becomes secondary		
⁄‿		
CANcel	⇒	CANcell**A**tion
iambs (weak **STRONG**)	+	-ation
Primary stress bumps left; English does not allow back-to-back primary stress		
‿⁄		
con**FIRM**	⇒	CONfirm**A**tion

Next teach derivational prefixes; these will add to the syllable count, but not affect primary stress assignment. Consider the following derived words and their roots:

3-2: immoral; dishonest; illegal; disloyal (derived from 2-1 roots)
3-3: immature; reinvent (derived from 2-2 roots)
5-3: unreliable (derived from a 4-2 root)

Sample sentence:

reliable (4-2) + prefix "un" ⇒ unreliable (5-3)
My boss is unreliable.

Now introduce contrastive stress:

unreliable (5-3) + Contrastive Stress ⇒ *un*reliable (5-1)
My boss used to be reliable, but lately she's become *un*reliable.

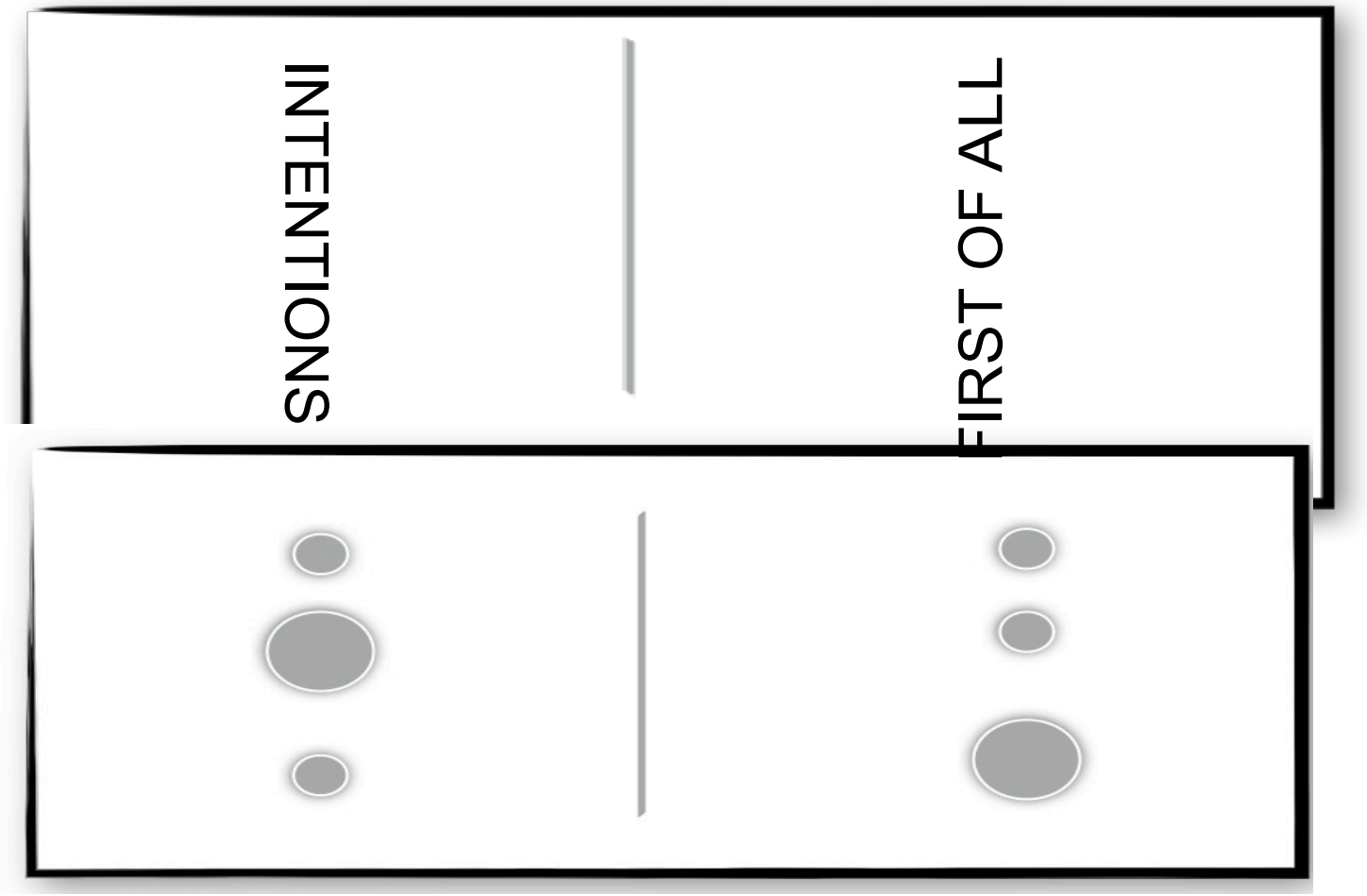
Challenges

Mastering the speech rhythm of English is reported to be difficult for learners (Wennerstrom, 2001). In addition to insufficient durational differences between stressed and unstressed syllables in words (Flege & Bohn, 1989), is the tendency not to reduce the duration of function words in sentences (Aoyama & Guion, 2007). The communication breakdowns that result can be at times humorous, as when the directive, "[θri tu faiv]" is intended as an indication of the hours of service [θri tɘ faiv], not the location or room number of the service (Grant, 2009, SUPRAS communication). While the temporal difference native speakers use to distinguish between "two" and "to" can be innocuous in one situation, the lack of durational differentiation between these homophones has proven to be fatal in aviation communication. This was the case in a 1989 crash reported by Cushing (1995), in which, as the clearance read back makes evident, the pilot understood the Air Traffic Controller (ATC) to be clearing him to descend to (preposition) "four zero zero" hundred feet. The aircraft crashed into a hillside at 437 feet, just before reaching the vector beacon, where the minimal descent height was, as the ATC intended to convey, 2400 feet.

> 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). (Cushing, 1995, p. 3)

Teaching tips Dominoes

Provide Practice Opportunities for Phrasal Stress. Revive and Revise the Word Dominoes. Create Word Dominoes that include preposition phrases and phrases with fixed stress patterns. Such phrases pattern just like lexical items. Use them to expand learner metacognition to the phrase level, and to promote practice with normal stress patterns in both multisyllabic words and in phrases.



Be sure to include in your Word Dominoes lexical items with stress patterns to match the phrasal patterns, both prepositional phrases and fixed phrases.

Sample phrasal stress patterns:

Prepositional Phrases		Fixed Phrases		Sample Words
2-2	4-3			
at work	in the morning	In other words…	4-2	affordable
at home	at the office	In the first place…	4-3	disagreement
at school	on the table	In the meantime…	4-3	combination
on time	for the weekend	On the other hand…	5-3	understandable
for now	to the movies	As I was saying…	5-4	characteristic

Provide Practice Opportunities for Contrastive Phrasal Stress in marked utterances.

> I said to pick me up *before* work, not *after* work.
> Although I didn't arrive *on* time, I made it *in* time, since everyone else was late, too.

Elicit Student Tell Backs: Raise student metacognition of the functions of contrastive stress.

> "We use Extra Stress on Function Words to show contrast."

Important news

Investigations of the stress patterns of non-native speakers at various levels of proficiency reveal relatively restricted pitch range and uniform stress; that is, all words stressed more or less equally regardless of lexical category ('content' versus 'function' words) or semantic weight (Wennerstrom, 2000; Tyler, Jeffries, & Davies, 1988). In her study of listeners' judgments on L2 accentedness and comprehensibility, Kang (2010) found that the low-proficiency International Teaching Assistants (ITAs) "placed stress on many function words or articles such as 'be', 'the', 'that', and 'this is' (p. 310) with the result that "the more stressed words ITAs produced, the more accented their speech sounded to undergraduate students" (p. 309).

Teaching tips

Introduce normal, unmarked sentence level stress with a minimalist matrix. Content words appear in base, citation form, minus inflectional morphology. Provide a Template Sentence to establish context and grammatical tense. Students generate sentences supplying obligatory function words and morphology. This matrix is versatile, providing opportunities for practice of Yes/No and WH-questions as well as practice with the normal sentence stress.
Sample Matrix, using the long-running Law & Order series, and accompanying directions:

defense attorney	**defend**	**client**
jury	**reach**	**verdict**
judge	**sentence**	**criminal**

Directions: Finish this sentence: In a typical trial, _____.
Example: In a typical trial, the defense attorney makes an opening argument.

Elicit Student Tell Backs: Extend metacognition from phrases to sentences.

> "In normal sentences, content words are stressed; function words are unstressed."

Introduce marked, implicational sentence stress with passages containing italicized words. Native speakers recognize that words in italics must be read differently. In the absence of instruction, learners

of English may be unaware of the use of italics as a deliberate or conscious signaling device used by the author. Learners may also be unaware of the oral equivalent of the printed italics, accounting for their failure to produce implicational fall/rise intonation.

Sample Sentence. Kenyan author Ngugi (1986, p. 11), writing about the effect of colonization on his native language:

"English became more than a language: it was *the* language, and all the others had to bow before it."

Extend instruction to marked, implicational stress in sentences with no equivalent use of italics in print. Create two sounds files and present the sentences orally only:

(Sentence 1) The teacher didn't grade your papers.
(Sentence 2) The *teacher* didn't grade your papers.

Introduce the Implicational Stress Checklist: co-construct the language to justify learner responses. Implicational Stress Checklist:

Do Sentences (1) and (2) sound The Same or **Different**
Explain your Choice: (e.g., "Sentence 2 was extra sing-songy/ used extra stress/...")

Student Tell Backs for Aural Discrimination:

"The vowel in the stressed syllable in the focus word was longer, louder, higher, clearer."

Extend the Implicational Stress Checklist: include the pragmatic function of implicational stress

Ask: "Have the papers been graded?" **YES** NO
Explain your Choice: (e.g., "Sentence 2 used extra stress on 'teacher' to signal an implication.")
State the Implication: Someone else (not the teacher) graded the papers.

Extend the Implicational Stress Checklist: introduce Given & New versus contrastive information.

Given & New: use normal sentence intonation:

Yesterday we discussed X. Today we'll discuss Y.
Yesterday we discussed the creation of MySpace. Today we'll discuss the creation of FaceBook.

Contrastive information: use contrastive stress to signal contrasting elements

Yesterday we discussed something about X. Today we'll discuss something else about X.
Yesterday we discussed the creation of MySpace. Today we'll discuss the marketing of MySpace.

Assess Listening Skills and Metacognition. Create two sound files; present orally only:

(1) There are some high-tech companies that sell a wide variety of products.

(2) There are *some* high tech companies that sell a wide variety of products.

Implicational Stress Checklist:

Do Sentences (1) and (2) sound the Same or **Different?**

Explain your Choice: (e.g., "Sentence 2 used extra stress on function word "some" to signal an implication.")

Assess metacognition of the pragmatic function of implicational stress

Ask: What will the speaker discuss next?

(a) the wide variety of products
(b) other high-tech companies

Choice (b) because the speaker used implicational stress on the function word "some".

Summary

The prosody of English begins with the alternation of stressed and unstressed syllables in disyllabic and polysyllabic words. The complexity of the English syllable poses challenges. To complicate matters, learners may hear as "exaggerated" what are actually the normal stress alternations of English, they may harbor negative views about implementing these in their own speech, and they may fail to conceptually grasp the intentional uses of marked, contrastive or implicational stress patterns by native speakers. As noted by Paunović & Savić (2008),

> "Students often do not have a clear idea of why exactly 'the melody of speech' should be important for communication, and therefore seem to lack the motivation to master it, while teachers do not seem to be theoretically or practically well-equipped to explain and illustrate its significance" (pp. 72-73).

These Teaching Tips are intended to facilitate successful acquisition of the prosody of English. The strategies and checklists are suggested in order to address metacognition and aural and oral skills. Target measurable learner outcomes include:

- monosyllabic words produced as monosyllables
- disyllabic and multisyllabic words produced with the correct primary stress
- aural detection and oral production of marked, contrastive or implicational stress

• metacognitive articulation of the pragmatic functions of marked, contrastive or implicational stress

ABOUT THE AUTHOR

Marnie Reed is Associate Professor of Education and affiliated faculty in the Program in Applied Linguistics at Boston University. She teaches courses in linguistics, second language acquisition, and second language phonology. She is co-author of the forthcoming second edition of *Sound Concepts: Strategies for Listening and Speaking* (University of Michigan Press), and co-editor (with John Levis) of the forthcoming *Handbook of English Pronunciation* (Wiley Blackwell).

REFERENCES

Aoyama, K., & Guion, S. G. (2007). Prosody in second language acquisition: Acoustic analyses of duration and F0 range. In Bohn, O-S., & Munro, M. J. (Eds*.), Language experience in second language speech learning: In honor of James Emil Flege* (pp. 282-297). Amsterdam: John Benjamins.

Archibald, John. (1993). Metrical phonology and the acquisition of L2 stress. In F. Eckman (Ed.), *Confluence: Linguistics, L2 Acquisition, and Speech Pathology*, pp. 37-48. Amsterdam, The Netherlands: John Benjamins Publishing Co.

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

Celce-Murcia, M., Brinton, D., Goodwin, J., & Griner, B. (2010). *Teaching Pronunciation: A Course Book and Reference Guide (2nd edition)*. New York: Cambridge University Press.

Coxhead, A. (2000). A new Academic Word List. *TESOL Quarterly, 34*(2), 213-238.

Cutler, A. (1986). Forebear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech, 29*(3), 201-220.

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.

Cutler, A., Dahan, D. & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40* (2), 141-20

Derwing, T. & M. Rossiter. 2002. ESL Learners' Perceptions of Their Pronunciation Needs and Strategies. *System*, 30 (2), 155-166.

Flege, J. E, & Bohn, O. S. (1989). An instrumental study of vowel reduction and stress placement in Spanish accented English, *Studies in Second Language Acquisition, 11*, 35–62.

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

Guion, S.G., Harada, T., & Clark, J.J. (2004). Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism Language and cognition*, *7*(3), 207-226.

Hancock, M. (1995). *Pronunciation Games*. Cambridge University Press.

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

Maddieson, I. (2013). Syllable Structure. In Dryer, M. S. & M. Haspelmath (Eds.) *The World Atlas of Language Structures*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/12

Murphy, J., Kandil, M. (2004). Word-level stress patterns in the Academic Word List. *System, 32* (1), 61-74.

Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge, UK: Cambridge University Press.

Ngugi, W-T. (1986). *Decolonizing the Mind The politics of Language in African Literature*. Portsmouth, NH: Heineman.

Pater, J. (1997). Metrical parameter missetting in second language acquisition. In S. J. Hannahs & M. Young-Scholten (Eds.), *Focus on phonological acquisition* (pp. 235–262). Amsterdam: John Benajmins.

Peperkamp, S., Vandelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics, 38,* 422-430.

Reed, M., & Michaud, C. (2005). *Sound Concepts: An Integrated Pronunciation Course*. NY: McGraw-Hill.

Reed. M. (2012). The effect of metacognitive feedback on second language morpho-phonology. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference*, Sept. 2001. (168-177). Ames, IA: Iowa State University.

Roach, P. (2009). *English phonetics and phonology, 4th edition*. Cambridge, UK: Cambridge University Press, p. 73.

Sluijter, A. M.C., van Heuven, V. J. (1996). Acoustic correlates of linguistic stress and accent in Dutch and American English. *Journal of the Acoustical Society of America, 100*, 2471-2485.

Tyler, A., Jeffries, A., & Davies, C. (1988). The effect of discourse structuring devices on listener perceptions of coherence in non-native university teachers' spoken discourse. *World Englishes, 7*, 101-110.

Vandergrift, L. & Goh, C. (2012). *Teaching and Learning Second Language Listening: Metacognition in Action.* NY: Routledge.

Wayland, R., Landfair, D., Li, B., & Guion, S.G. (2006). Native Thai speakers' acquisition of English word stress patterns. *Journal of Psycholinguist Research, 35*(3), 285-304.

Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggenback (Ed.), *Perspectives on Fluency*. Ann Arbor: University of Michigan Press.

Wennerstrom, A. (2001). *The music of everyday speech*. Oxford, England: Oxford University Press.

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