PROCEEDINGS OF THE 9TH ANNUAL PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING CONFERENCE

(ISSN 2380-9566)

Editor

John Levis

2018, Iowa State University
http://apling.engl.iastate.edu

TABLE OF CONTENTS

9TH PROCEEDINGS OF THE PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING CONFERENCE

Teachers & Learners

Variation

Instruction

Teaching Tips and Resources

Levis, J. M. (2018). PSLLT: Bridging L2 pronunciation research and practice. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. i-xii). Ames, IA: Iowa State University.

PSLLT:
BRIDGING L2 PRONUNCIATION RESEARCH AND TEACHING

John Levis, Iowa State University

The 9th Pronunciation in Second Language Learning and Teaching Conference, held in Salt Lake City, Utah, at the University of Utah, reflected developments in the field that have been evident for some years now. The most obvious development has been the connection of high-quality research conference with strong and vibrant connections to pedagogical questions and needs. The reason for this unusual connection is that pronunciation research is deeply connected to pronunciation teaching, and pronunciation teachers want to know what kinds of practices are evidence-based.

A second development that became obvious was explicit attention to research methodologies. Around 120 conference participants took part in the pre-conference Research Workshops that were generously funded by the University of Utah Vice President for Research. During these workshops, we learned from a variety of scholars about carrying out longitudinal research, increasing reproducibility in research, mixed effects modeling in pronunciation research, corpus linguistics and pronunciation, using ultrasound to visualize speech, and how to use Mechanical Turk for research.

Though not a development except in scope, and ongoing, valuable part of the conference for those who attend is the opportunities to network and interact with scholars one might otherwise not talk to. This can happen at any conference, of course, but the amount of interaction at PSLLT has always been unusual. Established scholars can be seen talking to up-and-coming researchers, those working in English and other visible languages talk to an help those who work in less commonly taught languages, those working in SL and FL contexts hear what the others are doing, those working with low-level learners provide insights to those working with advanced learners, and those who do very little research are able to talk to those doing lots of research, and vice versa. From all kinds of L2 pronunciation contexts, researchers and teachers, graduate students and faculty, those from North American and those from 10-25 countries around the world make connections. A professor, attending for his first time, told me that what most impressed him was that most conferences he goes to have a majority of established faculty members, but that this conference had a majority of younger faculty and graduate students. To me, this means the future is very bright. As established scholars come closer to retirement, the field is becoming younger and more diverse. Pronunciation is no longer an ESL/EFL/ELF question, but is a quickly growing area of interest in foreign language teaching as well, and is closely tied to changes in technology.

PSLLT (pronounced either “P-S-L-L-T” or “Pissalt” with stress on the first syllable) has now been away from its beginnings at Iowa State University five times in nine years, and has become a conference of choice for many researchers and teachers. This conference had 180 participants from all over the world, about 20% higher than previous numbers.

The program included six pre-conference research workshops, a new feature of the conference. The conference program is reproduced below, both as a matter of historical record and to help understand the wider context for the proceedings. After 9 years, we now have well over 250 articles in the proceedings. They reflect the range of paper that are presented each year, with many empirical studies, many teaching oriented papers, and from different languages.

The presentations in bold are those that are included in the proceedings. All presenters are invited to submit to the proceedings after the conference, but a minority ultimately submit their work. Others submit to journals in the field or get busy. All papers go through peer review, and those who revise in accordance with the reviewers' critiques are usually included in the proceedings.

Proceedings reviewers: Isabelle Darcy, Nadine De Moras, Tracey Derwing, Rachel Hayes-Harb, Amanda Huensch, Okim Kang, Edna Lima, Ivana Lučić, Shannon McCrocklin, Collen Meyer, Jessica Miller, Joan Mora, Charles Nagle, Mary Grantham O'Brien, Lucy Pickering, Marnie Reed, Alif Silpachai, Ala Simonchyk, Jessica Sturm, Ron Thomson, Lara Wallace, Germán Zaráte-Sández, Beth Zielinski.

PRONUNCIATION IN SECOND LANGUAGE LEARNING AND TEACHING (PSLLT) 2017

Bridging L2 Pronunciation Research and Teaching

SEPTEMBER 1-2, 2017

RESEARCH METHODS IN SECOND LANGUAGE PRONUNCIATION WORKSHOP

AUGUST 31, 2017

Image courtesy of the University of Utah

Dear Colleagues,

We are delighted to welcome you to Salt Lake City for Pronunciation in Second Language Learning and Teaching 2017. It is an honor to host this event, and we hope your time at the University of Utah is filled with stimulating and productive experiences.

Please do not hesitate to let any of us know if you need assistance and/or if there are ways we can help to make your time here enjoyable and successful.

Sincerely,

The 2017 PSLLT Organizing Committee

Rachel Hayes-Harb, Shannon Barrios, Catherine E. Showalter & Taylor Anne Barriuso

WORKSHOP AND CONFERENCE SPONSORS

- **Office of the Vice President for Research at the University of Utah**
- **Department of Linguistics at the University of Utah**
- **Second Language Teaching and Research Center (L2TReC)**
- Book donations from: John Benjamins, Routledge, and de Gruyter Mouton.

DEPARTMENT OF LINGUISTICS STAFF

Jessica Darrington
Shantel de Arraiz

PSLLT 2017 CONFERENCE VOLUNTEERS

Alisa Bedrov
Jenia Ivanova
Miranda McCarvel
Joselyn Rodriguez
Vicki Wason

PSLLT 2017 ABSTRACT REVIEW COMMITTEE

Shannon Becker
Tracey Derwing
Jenn Foote
Mara Haslam
Luke Harding
Amanda Heunsch
Talia Isaacs
Okim Kang
Sara Kennedy
John Levis
Alison McGregor
Murray Munro
Mary Grantham O'Brien
Lucy Pickering
Carolyn Pytlyk
Kazuya Saito
Mari Sakai
Ala Simonchyk
Sinem Sonsaat
Jessica Strum
Ron Thomson
Pavel Trofimovich
Elizabeth Zetterholm
Beth Zielinski

Research Methods in L2 Pronunciation Workshop		
THURSDAY, AUGUST 31, 2017		
11:30AM -	Speech Acquisition Lab Open House & Lunch - Lingusitics Department, LNCO 2300	
11:00AM -	Check-in & On-site Registration - LNCO 2945	
1:00PM-2:00PM	Workshop Session I	
	CTIHB JewelBox	*LNCO 2110*
	Practical steps for increasing openness and reproducibility in scientific research (Kidwell)	fears about longitudinal pronunciation research: Anxiety reduction through planning (Derwing, Munro)
2:30PM-3:30PM	Workshop Session II	
	CTIHB JewelBox	*LNCO 2110*
	An introduction to fitting and evaluating mixed-effects models in R	Corpus Linguistics and Pronunciation Analysis (Staples, Kang)
3:30PM-4:00PM	Snacks - Linguistics Department, LNCO 2300	
4:00PM-5:00PM	Workshop Session III	
	CTIHB JewelBox	*LNCO 2110*
	Visualizing speech in a classroom setting using interactive ultrasound	Introduction to using Mechanical Turk for linguistics research

PSLLT 2017 Conference

THURSDAY, AUGUST 31, 2017				
5:30PM-8:00PM	Check-in & on-site registration; Welcome Reception - Douglas Ballroom West + Patio			
FRIDAY, SEPTEMBER 1, 2017				
8:45AM-10:00AM	Welcome & Keynote - Douglas Ballroom, University Guest House			
10:00AM-	Coffee Break - Outside Douglas Ballroom West, University Guest House			
9:45AM-5:00PM	Check-in & On-site Registration - East Lobby, Officers' Club			
10:30AM-	Oral Session I			
	Officers' Club North Room (Session	*Officers' Club South Room*	*Officers' Club East Room (Session chair: Taylor Anne*	*Officers' Club West*
10:30AM-10:55AM	The role of L1-to-L2 sound matching relationship and orthographic information	Addressing ethical issues in classroom pronunciation assessment (Foote,	**The non-pronunciation teacher's guide to ESL pronunciation teaching (Cox, Henrichsen)**	Golden Speaker Builder: An interactive online tool for L2 learners to
11:00AM-11:25AM	Relationship between utterance fluency and cognitive fluency in first and	**Second language comprehensibility ratings: Do ESL and EFL teachers**	Corpus examination of pronunciation features identified in oral instructors' reports (Crabtree, Staples)	Non-native speech and recognition accuracy of two ASR
11:30AM-11:55AM	The vicissitudes of intelligible segmental	Accent and comprehensibility of English- speaking	From IPA to IPA: Bridging L2 pronunciation research and teaching	Improving novice learners' L2
12:00PM-12:25PM	**"Was that a question?" Perception of utterance- final intonation among L2 learners**	The effect of listener proficiency on comprehension and	Crossing the pronunciation bridge from isolation to integration	**Pronunciation learning through L2-**
12:30PM-2:00PM	Poster Session & Lunch (provided) - Douglas Ballroom, University Guest House			
2:00PM-3:25PM	Oral Presentation Session II			
	Officers' Club North Room (Session	*Officers' Club South Room*	*Officers' Club East Room (Session chair: Donna Brinton)*	*Officers' Club West*
2:00PM-2:25PM	Asymmetric processing of English Intonation: An eye-	The effectiveness of explicit pedagogical intervention in the L2	New methods for predicting perceptual discriminability of non-native contrasts (Daidone, Lidster,	**Focus on phonetic form and cognitive complexity in the**
2:30PM-2:55PM	The influence of language experience on L2	Tweeting about French pronunciation: Phonetic	L2 Japanese vowel production: A closer look at transfer effects from perception training (Hardison, Okuno)	Familiarity and congruence orthographic input
3:00PM-3:25PM	**It's about time: A longitudinal study on the development of L2 oral**	Teacher cognition of English intonation: Case study analysis of knowledge, beliefs, and thinking that	**Processing indexical and dialectal variation in a second language: Implications for the L2 classroom**	Accurate production in L2: A look behind the scenes
3:25PM-3:45PM	Coffee Break - East and West Lobbies, Officers' Club			
3:45PM-5:10PM	Oral Presentation Session III			
	Officers' Club North Room (Session	*Officers' Club South Room*	*Officers' Club East Room (Session chair: Lauren Brocious)*	*Officers' Club West*
3:45PM-4:10PM	**The role of duration in Japanese speakers' productions of English**	**Training the production of English L2 vowel contrasts through**	An investigation into pronunciation teaching practices in tertiary EFL classrooms in Vietnam (Nguyen,	**Student perceptions of university instructor accent in**
4:15PM-4:40PM	Accommodating different interlocutors: Nonnative speakers' use of phonetic alignment strategies	Types and sources of knowledge about pronunciation	**Accentedness and acceptability of Japanese English teachers' pronunciation: Ratings by three listener groups (Sugimoto, Uchida)**	Help or hinder? The complicated role of pronunciation
4:45PM-5:10PM	EFL learners' pronunciation performance in speech tasks with different levels of	**Traditional classroom pronunciation instruction: A**	EFL teachers' diagnostic assessment and pedagogical decisions in the teaching of pronunciation	Fluency and comprehensibility of 1st year and continuing graduate
7:00PM-10:00PM	Conference Dinner - Douglas Ballroom, University Guest House			

SATURDAY, SEPTEMBER 2, 2017				
9:00AM-	Check-in & On-site Registration - East Lobby, Officers' Club			
9:00AM-10:00AM	Publishing in Second Language Pronunciation (*Journal of Second Language Pronunciation* Editor John Levis) - Officers' Club South Room			
	Corpora in L2 Pronunciation Research: Preliminary Discussion on Ways Forward (Facilitated by Tracey Derwing and Mary O'Brien) - Officers' Club North Room			
10:00AM-11:55AM	Oral Presentation Session IV			
	Officers' Club North Room (Session chair:	*Officers' Club South Room (Session chair:*	*Officers' Club East Room (Session chair: Joshua*	*Officers' Club West Room*
10:00AM-10:25AM	The effect of vowel length on English as a Lingua Franca (ELF)	Teaching articulatory strategies: The role of	Improving teacher expertise in pronunciation instruction: A study circle model (Echelberger,	The search for an ideal scale length: What we can learn through
10:30AM-10:55AM	The role of suprasegmental features in L2 listeners' judgment of L2 English: A qualitative approach (Huang)	The development of L2 French learners' pronunciation, fluency, and comprehensibility: An online classroom study (Inceoglu)	**Pronunciation instruction practices of teachers of languages other than English (Huensch)**	Bilingualism in the Peruvian Amazon: Intervocalic stops in Yagua-Spanish and Bora- Spanish
11:00AM-11:25AM	Examining L1 effects in L2 German lexical stress assignment (O'Brien)	"This is how a gondolier gallops": Pronunciation and unintelligibility in ITA presentations (Levis, Muller Levis)	**Translanguaging in prosody teaching: Beyond monolingual ideologies (Liu)**	Self-evaluations in the acquisition of pronunciation of French as an L2
11:25AM-11:55AM	Listener perception of pronunciation and length of speech stimuli: Does length	Learners' attitudes towards visual feedback in pronunciation learning	**The impact of explicit instruction on the pronunciation of French**	A role for acoustic analysis in
12:00PM-	Lunch (on own) / JSLP Editorial Board Meeting			
2:00PM-	Teaching Tips Session A - Officers' Club South Room			
	Table 1 The use of MRI and ultrasound technology in teaching about Spanish (and general) phonetics and pronunciation (Holt) Table 2 A new comprehensive assessment tool for English pronunciation (Haslam) Table 3 Connecting the dots to L2 proficiency with an assessment template (Miller) Table 4 Integrating pronunciation, speaking, and listening through popular media in ESL classes (Arshavskaya) Table 5 Classroom mixers for pronunciation and listening (Chan) Table 6 The vowel elevator: A visual-kinesthetic way to			
2:55PM-3:45PM	Teaching Tips Session B - Officers' Club South Room			
	Table 1 From Broadway to the classroom: Using rap, prose and poetry for pronunciation (Martinez) **Table 2 The Tic Tac trick to teach the American English articulatory setting (McGregor)** Table 3 Stop shouting at me! (Meyers) Table 4 Meaningful feedback on pronunciation: Sneaking around the affective filter (Littlepage)			
3:45PM-4:00PM	Conference Closing - Officers' Club South Room			

v

LIST OF POSTERS

#1 French and Spanish pronunciation in CALL Software: Rosetta Stone, Duolingo, Babbel, and Mango Languages (Bajorek)
#2 The efficacy of high variability phonetic training in a non-laboratory setting (Barriuso)
#3 The frustrating case of French nasal vowels: Why our students confuse them and what we can do about it (Becker)
#4 Helping Vietnamese speakers acquire a listener-friendly pronunciation in English (Chan)
#5 Challenging the notion that Japanese English language learners cannot distinguish the /r/ and /l/ Phonemes (Chase, Tanner, Nissen, Hartshorn)
#6 **Empowering adult ELLs' fluency and pronunciation skills through readers theater (Chugg, Tanner)**
#7 Linguistic dimensions of L2 accentedness and comprehensibility vary across speaking tasks (Crowther, Trofimovich, Saito, Isaacs)
#8 **Learners' perspectives on English pronunciation teaching and learning: A preliminary study in the Vietnamese context (Dao)**
#9 Task-based assessment of academic English pronunciation (Domby)
#10 **Self-correction of second-language pronunciation via online, real-time, visual feedback (Garcia, Kolat, Morgan)**
#11 A snapshot of native and non-native Spanish vowel production across word boundaries (Henderson, Willis, Holt)
#12 **Online resources for learners and teachers of English pronunciation (Henrichsen)**
#13 Prosodic constructions in English dialog: Form, function, learner needs, and teachability (Ward)
#14 An intelligibility-based approach to English vowel pronunciation teaching in Korean context (Kang, Ahn)
#15 Integrating pronunciation teaching into the curriculum for beginner-level adult migrants (Keenan, Corrigan)
#16 Language input and the acquisition of Japanese lexical rhythm (Kinoshita, Sheppard)
#17 The effect of interlanguage speech intelligibility and attitudal benefit on speech perception (Koo)
#18 Pronunciationforteachers.com – A resource for pronunciation teaching and research (Levis, Sonsaat)
#19 ESL learners' experiences using electropalatographic biofeedback to improve pronunciation (Li, Tanner, Nissen, Hartshorn)
#20 Native listeners' assessment of L2 speech comprehensibility: What features matter most in North American English? (Lima)
#21 A comparative study of English and Mandarin discourse prosody (Liu)
#22 Intelligibility of Japanese-accented pronunciation of English: A phonetic analysis based on English read by Japanese database (Makino)
#23 Online oral practice platform Speak Everywhere for daily pronunciation practice (Miyamoto, Suzuki, Fukada)
#24 The role of prosody in signaling rhetorical organization (Mohammed)
#25 The big, bad [ə] (Mroz)
#26 Acquisition of L2 Japanese rhythm: How does durational variability change over time? (Mukai, Aoki, Daiju)
~~#27 L2 English production and perception by L1 Tera speakers: The effect of instruction (Musa)~~
#28 Segment or feature acquisition?: Generalizability of phonetic gains in L2 production (Olson)
#29 Promoting naturalistic L2 language acquisition through digitized contrastive speech training (Penssler-Beyer)
#30 **Stereotypical accent and French pronunciation learning (Ruellot)**
#31 Study abroad benefits in the comprehension of dialectal speech (Schmidt)
#32 The perception of lexical tones by native speakers of Korean with and without Mandarin learning experience (Tsukada, Han)
#33 High variability phonetic training and L2 lexical tones (Silpachai)
#34 **Prominence and information structure in pronunciation teaching materials (Silpachai, Levis)**
#35 **Preservice English teachers' perspectives on learning and teaching pronunciation in Turkey (Uzun, Ay)**
#36 English intonation produced by EFL Spanish speakers before and after Praat training (Valenzuela Farias)
#37 6 Ways to Use YouGlish to increase input and focus pronunciation practice (Wallace)

RESEARCH METHODS IN L2 PRONUNCIATION RESEARCH

OVERCOMING YOUR FEARS ABOUT LONGITUDINAL PRONUNCIATION RESEARCH: ANXIETY REDUCTION THROUGH PLANNING

Presented by Tracey Derwing & Murray Munro

Description

We will draw on our own experiences with a ten-year, largely quantitative study (Derwing & Munro, 2015) and other recent longitudinal work. Participants will consider a research question regarding an aspect of pronunciation learning, and together we will go through the stages necessary to design a longitudinal study. Target sample size and attrition, developing relationships with programs and participants, options for data analysis, and devising tangential studies will all be discussed. Together we will develop a checklist of strategies to guide prospective researchers toward a methodical and practical approach to conducting longitudinal research.

Learning outcomes

To conduct effective longitudinal research, it is important to try to anticipate problems that can arise and to mitigate their effects. Of particular concern are such challenges as a closed subject set, attrition, test focus, testing effects, missing data, and demotivation of participants. Without considering these issues in advance, researchers tread in risky waters indeed. In this workshop, participants will gain an understanding of how to implement the core aspects of a longitudinal study through (a) extensive planning prior to initiation of a project; (b) creating designs that allow for incremental publication, rather than waiting until the end; (c) incorporating flexibility to recognize opportune but previously unanticipated studies; (d) teamwork; and (e) organization and management of data, taking into account changing technological platforms.

About the presenters

Tracey Derwing is Professor Emerita at the University of Alberta and an Adjunct Professor in Linguistics at Simon Fraser University. Her primary interests are factors affecting communicative success for L2 speakers. Murray Munro is a professor of Linguistics at Simon Fraser University. He is an applied phonetician whose interests include L2 pronunciation and forensic linguistics.

PRACTICAL STEPS FOR INCREASING OPENNESS AND REPRODUCIBILITY IN SCIENTIFIC RESEARCH

Presented by Mallory Kidwell

Description

This practical workshop will review laboratory and personal research practices to improve reproducibility and create more rigorous, open, and impactful research. Topics covered in this

workshop will include defining reproducible research, understanding statistical and organizational barriers to reproducibility, and learning how to use the Open Science Framework (https://osf.io/) to manage data and supplementary materials, facilitate collaborations, contain bias, and extend the reach of your research through private or public laboratory operations. This workshop does not require any specialized knowledge to participate but it will be hands-on, so please bring a laptop.

Learning outcomes

By the end of this workshop, participants will be able to (1) distinguish between different types of reproducibility and their barriers, (2) understand the impact of power and bias on reproducibility, (3) set up a research project using the Open Science Framework (OSF), and (4) use the OSF to keep research more organized, reproducible, and extendable.

About the presenter

Mallory Kidwell is a graduate student in the Clinical Psychology department at the University of Utah. Previously, she held the position of Metascience Project Coordinator at the Center for Open Science, coordinating projects that empirically evaluated reproducibility and the efficacy of open science practices in published scientific literature. Presently, she studies the psychophysiological mechanisms of risky or resilient behavior among adolescents exposed to trauma.

AN INTRODUCTION TO FITTING AND EVALUATING MIXED-EFFECTS MODELS IN R

Presented by Charles Nagle

Description

Mixed-effects modeling is a multidimensional statistical analysis that allows the researcher to partition and explain sources of within- and between-subjects variation by systematically manipulating the fixed and random effects structures of the model. Mixed-effects models are advantageous over more widely employed ANOVA because they are robust in the face of missing data and do not impose the same restrictions related to independence of observations. This workshop concentrates on fitting mixed-effects models in RStudio using the lme4 package. Participants will be provided with a brief overview of mixed-effects models and two methodological review articles before the workshop.

Learning outcomes

By the end of this workshop, participants will be able to (1) describe the advantages of mixed-effects modeling,
(2) read data into RStudio and examine its structure, and (3) fit and evaluate basic mixed-effects models, including the unconditional and unconditional linear growth models, using the code provided. Participants will work with a longitudinal data set and an R script containing annotated code for all of the operations to be carried out during the workshop. A list of resources related to

mixed-effects models will also be provided.

About the presenter

Charles Nagle is an Assistant Professor of Spanish in the Department of World Languages and Cultures at Iowa State University. His research examines individual differences and their relationship to pronunciation development over time.

CORPUS LINGUISTICS AND PRONUNCIATION ANALYSIS

Presented by Shelley Staples and Okim Kang

Description

Corpora can help us to understand actual patterns of language use, including pronunciation, in particular contexts and by particular language learners. However, the use of corpora in the analysis of pronunciation features has been limited. At the same time, speech science has made progress toward identifying various pronunciation features that can show second language (L2) learners' progress of language learning or predict their proficiency. It has become common for elements of speech production to be detected by instrument and computer-assisted acoustical analysis (e.g., PRAAT), which characterizes different aspects of pronunciation by examining patterns of speech properties.

Learning outcomes

This workshop will provide an introduction to using corpora for examining pronunciation features, including 1) building corpora of speech samples; 2) using Praat with pronunciation analysis; 3) coding corpus texts for pronunciation features; 4) using a free program to explore the use of coded texts to examine patterns.

About the presenters

Shelley Staples is an Assistant Professor in the English Applied Linguistics Program at University of Arizona. Her research focuses on the use of corpus linguistics for applications to language teaching and language assessment. Okim Kang is an Associate Professor in the Applied Linguistics Program at Northern Arizona University. Her research concerns aspects of L2 pronunciation, speech perception and production, automated speech scoring, oral language proficiency assessment, and language attitudes.

INTRODUCTION TO USING MECHANICAL TURK FOR LINGUISTICS RESEARCH

Presented by Emily Moeng

Description

The goal of this workshop is to introduce researchers to Mechanical Turk, with a focus on behavioral experiments that may require auditory stimuli. The workshop is primarily aimed at those with little to no experience with Mechanical Turk. This workshop will also introduce researchers to JsPsych, a JavaScript library for creating web-run human behavior experiments, including those that might require auditory stimuli. JsPsych is particularly well-suited for those who wish to run behavioral experiments through MTurk. Although some web coding experience is suggested, neither web coding nor experience with JavaScript are required.

Learning outcomes

By the end of this workshop, participants will (1) be equipped with the basic tools and knowledge to run an experiment through Mechanical Turk, (2) have experience with using JsPsych to create web-based behavioral experiments, (3) have an understanding of the limitations of an experiment conducted online and how to best adapt procedures for such an experiment, and (4) have an awareness of some pros and cons on conducting linguistic experiments through Mechanical Turk.

About the presenter

Emily Moeng is a graduate student in the Linguistics Department at the University of North Carolina at Chapel Hill. She studies phonological acquisition, specifically how acquisition at one level of categorization affects categories at other levels.

VISUALIZING SPEECH IN A CLASSROOM SETTING USING INTERACTIVE ULTRASOUND IMAGING

Presented by Sonya Bird and Heather Bliss

Description

This workshop will introduce participants to ultrasound imaging as a tool for teaching and researching pronunciation. We will review technical details and practical issues around ultrasound imaging, including affordability and accessibility. The challenges and benefits of using ultrasound across various contexts will be discussed. We will provide concrete examples of paradigms that can be used concurrently to both teach pronunciation using ultrasound and research its effectiveness. The workshop will include an interactive ultrasound session, in which participants will have the chance to practice their articulation of specific sounds and sound sequences that can pose challenges for learners.

Learning outcomes

Learning outcomes (i) Participants will learn about the benefits and limitations of using ultrasound

imaging technology as a tool for researching and teaching pronunciation of challenging sounds and sound sequences. (ii) Participants will acquire practical information on issues such as accessibility and affordability of ultrasound, as well as examples of teaching and research methods for different learning settings (e.g., individual learners, small groups, classrooms). (iii) Participants will gain hands-on and practical experience in how to use ultrasound imaging through an interactive session in which they can practice their own articulations using an ultrasound machine. (iv) Participants will gain an appreciation of the value of visualizing speech in teaching and learning the sounds of unfamiliar languages.

About the presenters

Sonya Bird is an Associate Professor in the Department of Linguistics at the University of Victoria. She studies the phonetic structures of Indigenous languages of the Pacific Northwest. Heather Bliss is a Banting Postdoctoral Fellow in the Department of Linguistics at the University of Victoria and an Adjunct Professor in the Department of Linguistics at the University of British Columbia. Her work focuses on Indigenous language documentation.

KEYNOTE ADDRESS

Professor Isabelle Darcy, *Indiana University*

"A psycholinguist walks into a classroom…": A road-map for bridging research and practice in pronunciation teaching

Over the last 50 years, our knowledge of how learners acquire the phonology and the pronunciation of a foreign language has made tremendous progress. During the same period, pronunciation teaching has also profoundly transformed itself. Yet there is a general feeling that the two fields are disconnected: that research is not asking the right questions, and that teaching practices are not taking research outcomes into account.

In an attempt to enhance cross-pollination between the fields, I will synthetize psycholinguistic research findings obtained both in and outside my lab - outlining phonological acquisition in L2 learners. I will highlight their relevance for pronunciation teaching and outline concrete ways to implement specific practices suggested by these findings. This is only a first step however, and actual research verifying the success of such implementations in increasing intelligibility is needed. The talk will thus outline a road-map for a research-teaching double agenda.

Gordon, J. (2018). Second-language comprehensibility ratings: do ESL and EFL teachers rate in the same way? In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 1-16). Ames, IA: Iowa State University.

SECOND LANGUAGE COMPREHENSIBILITY RATINGS: DO ESL AND EFL TEACHERS RATE IN THE SAME WAY?

Joshua Gordon, University of Northern Iowa

This mixed-method investigation analyzed the way native-ESL and nonnative-EFL teachers rated second language (L2) comprehensibility in ESL learners. Two groups of native-ESL and nonnative-EFL teachers rated spontaneous speech samples from ESL learners obtained before and after a stand-alone pronunciation course in an intensive ESL program. Speech samples from a group of L1-English speakers were also included for control purposes. The quantitative analyses indicated that the group of nonnative-EFL teachers rated the speech samples more severely than the native-ESL teachers. Additionally, stimulated-recall interviews carried out with each teacher after the rating task revealed similarities and differences between both groups of teachers. However, the most important differences seemed to be rooted in teachers' pedagogical knowledge, teacher training, and familiarity (or lack of familiarity) with different L2 accents. The results of this investigation are discussed in terms of their implications for L2 pronunciation instruction and teacher training.

INTRODUCTION

Studies that have explored how native (NSs) and nonnative speakers (NNSs) rate second language (L2) speech for comprehensibility (i.e., the perceived ease or difficulty of understanding a message, see Derwing & Munro, 2009) have found mixed results. Whereas NNSs have been found to rate L2 speech more severely than NSs (Fayer & Krasinski, 1987; Kang, 2012; Rossiter, 2009), other studies have found the opposite (Brown, 1995), while other investigations have not found any significant differences between both groups of listeners (Crowther, Trofimovich, & Isaacs, 2016; Derwing & Munro, 2013; Flege, 1988; MacKay, Flege, & Imani, 2006). However, the ways that L2 teachers rate L2 comprehensibility has remained mostly unexamined. Such a question is important to understand since the teacher is, in the majority of cases, the one who makes pedagogical decisions in class as to which students need to work more on their pronunciation. The mixed-method investigation which is discussed here analyzed the way native-ESL and nonnative-EFL teachers rated comprehensibility in ESL learners in order to explore the differences and similarities in how teachers rate, and to understand the types of criteria used by both groups of teachers to rate L2 speech. In the next pages, I will present a review of literature and previous studies that motivated my investigation, followed by the details of the methodology used to carry out this study. Finally, I will present quantitative and qualitative analyses, followed by a discussion of the significance of these findings for pronunciation instruction and teacher education.

Literature Review

Studies that have investigated what makes L2 speech comprehensible have demonstrated that in addition to phonological issues, lexical and syntactical factors can also affect the degree of comprehensibility in L2 speech (Crowther et al., 2016; Derwing & Munro, 2015; Trofimovich & Isaacs, 2012). In contrast, other studies that have investigated what constitutes a foreign accent in

L2 speech have determined that this dimension is mostly linked to phonological factors (see Derwing & Munro, 2009, 2015). Whereas these studies have analyzed what constitutes comprehensibility and foreign accent in L2 speech, other studies have investigated whether NSs or NNSs perceive L2 speech in the same way in terms of comprehensibility; the results of such studies have presented mixed results.

A few studies have determined that NNSs have a tendency to be stricter when it comes to rating L2 speech. For example, Kang (2012) investigated the effects of raters' backgrounds when rating the L2 speech of international teaching assistants (ITAs) from American universities. In this study, 70 undergraduate students (48 NSs and 22 NNSs of English) rated speech samples of 11 prospective ITAs from different language backgrounds. Kang found that NNSs were more severe than the NSs in their ratings, which confirmed previous results in which NNSs were stricter than NSs in rating L2 speech (e.g., Fayer & Krasinski, 1987; Rossiter, 2009). In other studies, however, the NSs have appeared to be more severe when rating L2 speech. For instance, Brown (1995) investigated the effects of raters' occupational and linguistic backgrounds at the moment of assessing L2 speakers of Japanese. In this case, 33 NSs and NNSs of Japanese rated speech samples of 51 NNSs of Japanese. Brown found that the NNSs were more lenient in their ratings when compared to the NSs.

The difference between NSs' and NNSs' ratings of L2 speech is even more complex in the sense that, contrary to the above-mentioned studies, other studies have not found significant differences between these groups of raters. In these studies, this lack of difference between NS and NNS raters may be rooted in the L2 proficiency of the NNSs. For example, in studies in which NNS raters had very high levels of L2 proficiency, those NNS raters typically tended to provide rating scores that were not significantly different from those of the NS raters (e.g., Derwing & Munro, 2013).[1] In a recent study, Crowther, et al. (2016) investigated the effects of listener status (NS vs. NNS) and language background when rating comprehensibility and accentedness of L2 speech. For this study, 26 NNSs of English (speakers of L2 English with French and Mandarin L1 backgrounds) rated spontaneous speech samples of 40 L2-English speakers (whose L1 was French). The ratings were compared with those of 60 NSs of English. There were no significant differences in the global ratings between the 60 NSs and the 26 NNSs, or between the two groups of NNSs. These results also agreed with other studies that did not find significant differences in the L2 speech ratings of NSs and NNSs (e.g., Derwing & Munro, 1997; 2013; Flege, 1988; MacKay, et al., 2006).

In addition to differences between NSs and NNSs, other studies have pointed out that different types of raters can provide different results. For example, Kennedy and Trofimovich (2008) stated that expert raters, like language teachers or linguistics students, can provide biased ratings because they regularly hear different accents, which could lead to more lenient ratings than those of an inexperienced rater (see also Thompson, 1991). Additionally, other studies have argued that familiarity with a specific accent (e.g., because of knowledge of another language) can also produce biased results. For example, Winke and Gass (2013) found that raters' familiarity with speakers' L1 affected how they came up with ratings of specific L2 speech samples, as opposed to other raters who were not familiar with specific languages or accents. In other words, this is an

[1] I would like to thank Tracey Derwing for bringing this issue to my attention.

area that has presented inconclusive results, as sharing the same L1 background between speakers and raters has been an advantage in some studies (e.g., Bent & Bradlow, 2003; Harding, 2012), but it has not been a determining factor in enhancing comprehensibility in listeners familiar with certain L1s in other studies (see Munro, Derwing, & Morton, 2006).

Other studies have analyzed how speaker-independent factors affect L2-speech perception. Levi, Winters, and Pisoni (2007) examined the effects of listening context and lexical frequency in the perception of foreign-accented speech. Listeners rated foreign accent in individual words produced by L1 and L2 speakers of English. The words were presented to the listeners in two contexts: auditory-only, and auditory+orthography. The results demonstrated that listeners perceived high-frequency words as less accented. Additionally, the use of orthographic cues caused L1 speakers of English to be perceived as less accented, whereas L2 speakers were rated as more accented. Levi et al. suggested that the high frequency words were probably rated as less accented because the listeners had more stored exemplars of them—that is, their representations were more robust in memory, so it was easier to perceive them in the rating task.

In a similar line of research, Saito and Shintani (2016) investigated the effects of bilingualism in the perception of comprehensibility of L2 speech. In this study, spontaneous speech samples from 50 L2-English speakers (L1-Japanese speakers) were presented to two groups of raters: 10 Canadian monolingual L1-English speakers, and 10 Singaporean L1-English speakers who were also proficient in other languages. Their results demonstrated two interesting findings. First, ratings were mainly influenced by phonological and temporal cues regardless of the raters' backgrounds. Second, the Singaporean raters were more lenient in rating comprehensibility. Saito and Shintani suggested that the Singaporean listeners, as speakers of different languages, were better accustomed to different phonological patterns, which made them more accommodating than the monolingual speakers when parsing L2 speech.

These conflicting results—ranging from the different aspects that listeners pay attention to in L2 speech to how raters' different language backgrounds affect how they rate—demonstrate that the factors underlying ratings of L2 comprehensibility are complex. Certainly, this type of research is clearly important for L2 pronunciation teaching and teacher education purposes; however, one group of raters that remains mostly unexamined is language teachers, and more specifically, NS and NNS teachers and how they may rate differently in different teaching contexts (i.e., second and foreign language contexts). This is especially important in the classroom, as language teachers decide which students need to work more on their pronunciation. However, as mentioned earlier, teachers' perceptions of their students' speech can be biased (Kennedy & Trofimovich, 2008) and their familiarity with different accents, or their NS/NNS status, can affect L2 speech perception (see Brown, 1995; Crowther et al., 2016; Kang, 2012; Winke & Gass, 2013). Therefore, it is important to compare the comprehensibility ratings of groups of teachers in different contexts.

The Current Study

In this study, I investigated how native-ESL and nonnative-EFL teachers rated L2 speech samples for comprehensibility to determine whether there were differences or similarities in the teachers' ratings, and to analyze on which aspects these two groups of teachers based their ratings. The study was guided by the following research questions:

1. Do ESL and EFL teachers rate L2 speech comprehensibility in the same way?

2. On which linguistic aspects (i.e., phonological, lexical, grammatical) do ESL and EFL teachers base their ratings?

METHODOLOGY

Three steps were followed to carry out this study. First, speech samples were collected from L2 learners in two stand-alone pronunciation classes in an intensive ESL program at a large American university in the Midwest. Speech samples from L1-English speakers were also collected as a control. Second, two groups of teachers (native-ESL & nonnative-EFL teachers) rated the speech samples for comprehensibility. Finally, stimulated-recall interviews were carried out with each teacher after the rating task to discuss their reasons for their ratings.

Stimuli

The speech samples used in the rating task were part of a corpus obtained from a previous study (see Gordon Zamora, 2015). Speech samples in the form of video-description narratives were collected from 10 ESL learners in two stand-alone pronunciation classes (Class A & Class B hereafter) in an intensive ESL program. There were 5 ESL learners in each class, who were enrolled in the highest two institutional levels of the program. They came from several countries, and spoke different L1s (see Table 1). These learners recorded descriptions of two different video cartoons found on the internet (Simon's Cat, 2009, 2010) at the beginning of their course (Time 1) and also at the end of their course (Time 2).[2] In addition to recordings from these L2 learners, 5 L1-English speakers also recorded descriptions of both video cartoons. These L1-English speakers were undergraduate students from the same school, and they recorded descriptions of both video cartoons only once. Because of the spontaneous nature of the samples, all the narratives were first transcribed to find similarities in the descriptions. There was a specific event in the plot of each cartoon that all the ESL learners and the L1-English speakers described across the board. Thus, passages between 15 or 17 seconds of those descriptions were presented to the two groups of ESL and EFL teachers in a rating task.

[2] The video cartoons Simon's Cat (2009) and Simon's Cat (2010) were presented to the ESL learners in Class A and B in inverted order to maintain a balance. For instance, the video that learners in Class A watched and described at Time 1 is the same video that learners in Class B watched and described at Time 2. In the same way, the video that learners in Class B watched at Time 1 is the same video that learners in Class A watched at Time 2.

Table 1.

ESL Learners from Class A and Class B.

CLASS A

ESL Learner	Country of Origin	Age	Native Language	Level in the Program
CAP01	Japan	20	Japanese	Level 7
CAP02	South Korea	25	Korean	Level 6
CAP03	Argentina	46	Spanish	Level 6
CAP04	Japan	20	Japanese	Level 7
CAP05	Kazakhstan	21	Kazakh	Level 6

CLASS B

ESL Learner	Country of Origin	Age	Native Language	Level in the Program
CBP01	Japan	21	Japanese	Level 7
CBP02	Saudi Arabia	35	Arabic	Level 7
CBP03	Japan	19	Japanese	Level 7
CBP04	Chile	40	Spanish	Level 6
CBP05	South Korea	20	Korean	Level 7

The Raters

All the speech samples were presented to a group of 12 native-ESL teachers and a group of 15 nonnative-EFL teachers. The ESL teachers were all born in the U.S., and all were L1 speakers of American English. All were instructors in two intensive ESL programs at two universities in the Midwest; they included both novice and experienced teachers, and all had advanced degrees in TESOL, Linguistics, or Applied Linguistics. All of them spoke at least one L2, and the majority even spoke or at least had knowledge of a third language. Additionally, each ESL instructor had taken a course on pedagogical phonology (i.e., a specific course where they studied how to teach L2 pronunciation), and the majority had lived in a country where one of their L2s was spoken (see Table 2). The EFL teachers were all from Costa Rica and were L1 speakers of Spanish. They all taught at the university level and had advanced degrees in teaching EFL. The majority worked as EFL teacher trainers preparing preservice teachers, and others worked as EFL

instructors for students from other majors.[3] The EFL teachers had taken only basic linguistics and phonetics/phonology courses as part of their training. They spoke only one L2 (English), and very few had knowledge of a third language at a basic reading level. In addition, most of the EFL teachers had not spent a considerable amount of time living in a country where their L2 was spoken (see Table 3). These two groups of teachers carried out an individual rating task, described below.

Table 2

Native-ESL Teachers' Background.

ESL Teacher	Age	Gender	Second Language	Self-Rated Proficiency	Other Languages	Self-Rated Proficiency	Highest Academic Degree	Years of Teaching Experience	Time Living Abroad
ESLN01	26	F	Spanish	Intermediate	Turkish	Beginner	M.A.	3	Spain, 6 months
ESLN02	25	F	Serbian	Intermediate	Russian	Intermediate	M.A.	2	Bosnia, 8 months
ESLN03	24	F	Spanish	Advanced	French	Beginner	M.A.	3	NA
ESLN04	23	F	German	Intermediate	Swedish	Beginner	M.A.	3	NA
ESLN05	24	M	Japanese	Advanced	Chinese	Beginner	M.A.	3	NA
ESLS01	32	F	Spanish	Advanced	NA	NA	M.A.	5	Mexico, 5 months
ESLS02	60	F	Dutch	Intermediate	NA	NA	M.A.	9	Holland, 20 years
ESLS03	45	M	Japanese	Beginner	NA	NA	M.A.	11	Japan, 1 year
ESLS04	30	F	Kyrgyz	Intermediate	Spanish	Intermediate	M.A.	7	Kyrgyzstan, 2 years
ESLS05	38	F	Spanish	Advanced	Quechua	Beginner	M.A.	12	Panama, 2 years
ESLS06	48	M	Japanese	Advanced	Mandarin	Intermediate	M.A.	15	Japan, 3 years
ESLS07	30	M	Spanish	Advanced	French	Intermediate	M.A.	7	Spain, 2 years

[3] Although no specific test was given to any of the teachers to assess their L2 proficiency (and the majority of EFL teachers in fact self-rated their L2 proficiency as "advanced" or "near-native"), the group of EFL teachers represented a homogeneous group of proficient speakers who had high levels of their L2 in order to perform some of their teaching duties (e.g., teacher training preservice teachers in different areas).

Table 3

Nonnative-EFL Teachers' Background.

EFL Teacher	Age	Gender	Second Language	Self-Rated Proficiency	Other Languages	Self-Rated Proficiency	Highest Academic Degree	Years of Teaching Experience	Time Living Abroad
EFLN01	24	M	English	Advanced	NA	NA	Lic.	2	NA
EFLN02	26	F	English	Advanced	NA	NA	Lic.	2	NA
EFLN03	25	F	English	Advanced	NA	NA	Lic.	3	NA
EFLN04	25	F	English	Advanced	NA	NA	Lic.	4	NA
EFLN05	23	M	English	Advanced	French	Reading Knowledge	Lic.	1	NA
EFLS01	36	F	English	Advanced	NA	NA	M.A.	13	NA
EFLS02	36	M	English	Advanced	NA	NA	M.A.	12	U.S. 1 month
EFLS04	33	F	English	Advanced	NA	NA	M.A.	10	U.S. 4 months
EFLS05	35	F	English	Advanced	French	Reading Knowledge	Ms.C.	13	NA
EFLS06	45	F	English	Near-Native	French	Reading Knowledge	M.A.	17	NA
EFLS07	33	F	English	Advanced	NA	NA	M.A.	11	NA
EFLS08	36	F	English	Advanced	NA	NA	M.A.	14	NA
EFLS09	34	M	English	Advanced	NA	NA	M.A.	12	NA
EFLS10	33	M	English	Advanced	NA	NA	M.A.	12	NA
EFLS11	34	F	English	Advanced	NA	NA	M.A.	13	NA

Rating Task & Stimulated-Recall Interviews

The rating task was carried out individually on a personal computer in a quiet office or library. First, the teachers watched the two video cartoons described by the speakers to avoid biased ratings with the initial samples (see Derwing, Rossiter, Munro, & Thomson, 2004). They also carried out a short warm-up to get familiar with the task.[4] The teachers then completed the rating task, which was programmed using PRAAT (Boersma & Weenink, 2015). The teachers heard the randomized speech samples through high-quality headphones, and rated each one by clicking

[4] For the warm-up task, the raters listened to 10 different sentences spoken by individuals who did not participate in the actual task. The sentences did not have any relationship with the videos being described by the speakers in the actual task.

on a 9-point Likert scale on the computer screen (1=*extremely easy to understand*, 9=*impossible to understand*; see Munro & Derwing, 1995).

Once the rating task was over, ten samples were randomly selected for a stimulated-recall interview with each teacher (Gass & Mackey, 2000). The samples were played one by one, and the teachers were reminded of the ratings given to each specific sample. They were asked about possible problems that made the speech samples difficult to understand (see Appendix 1). Further questions were asked for clarification, and if the teachers requested to hear a sample again, it was played as many times as necessary. Finally, all the stimulated-recall interviews were transcribed and comments were classified according to common themes (see Richards, 2003).

RESULTS

Quantitative

The inter-rater reliability coefficients (Cronbach's alpha) computed across all ratings given by the ESL and EFL teachers were very high at both times 1 and 2 (ESL = T1 .95, T2 .95; EFL = T1 .97, T2 .97). This indicated a strong agreement (Larson-Hall, 2009). As expected, both groups of raters found the NS control group the most comprehensible, as seen in the mean scores reported in Table 4.

Table 4.

Mean Comprehensibility Ratings.

Rater Group	Group	Time 1	Std. Error	Time 2	Std. Error
ESL Teachers	Class A (*n*=5)	3.96	0.34	4.51	0.35
	Class B (*n*=5)	3.57	0.34	3.57	0.36
	NS Group (*n*=5)	1.10	0.34	1.12	0.36
Rater Group	**Group**	**Time 1**	**Std. Error**	**Time 2**	**Std. Error**
EFL Teachers	Class A (*n*=5)	4.99	0.33	5.54	0.34
	Class B (*n*=5)	5.12	0.33	4.64	0.34
	NS Group (*n*=5)	1.52	0.33	1.42	0.34

A mixed-method analysis was carried out to determine if there were significant differences in the ratings of both groups of teachers based on class or time. The Type 3 Test of Fixed Effects found a significant effect of *Time*, $F(1, 791) = 1.20$, $p = 0.2735$, and no interaction of *Rater Group* over *Time*, $F(1, 791) = 1.59$, $p = 0.2073$. These results further strengthen the consistency of both groups of teachers in their ratings at times 1 and 2. However, the results also yielded a significant effect of *Rater Group*, $F(1, 791) = 41.49$, $p < 0.0001$, according to which the EFL teachers consistently rated the three groups of speakers (Class A, Class B, and NSs) as less

comprehensible, compared to the ESL teachers. Figure 1 below shows the different ratings given by the two groups of teachers over time (in general), whereas Figure 2 shows the same ratings given by both groups of teachers to the three different speaker groups (Class A, Class B, and NS).

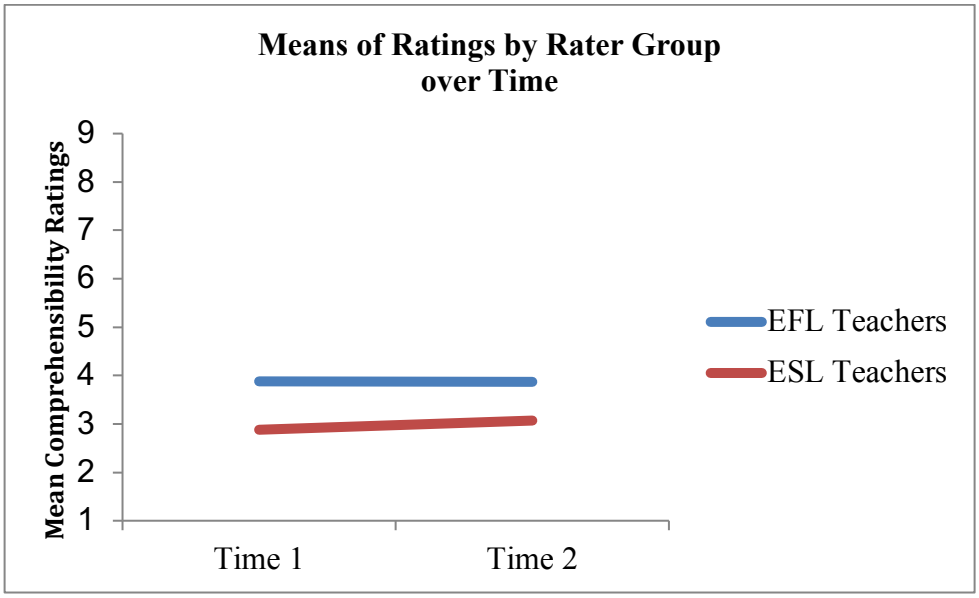


Figure 1. *Means of Ratings by Rater Group over Time.*

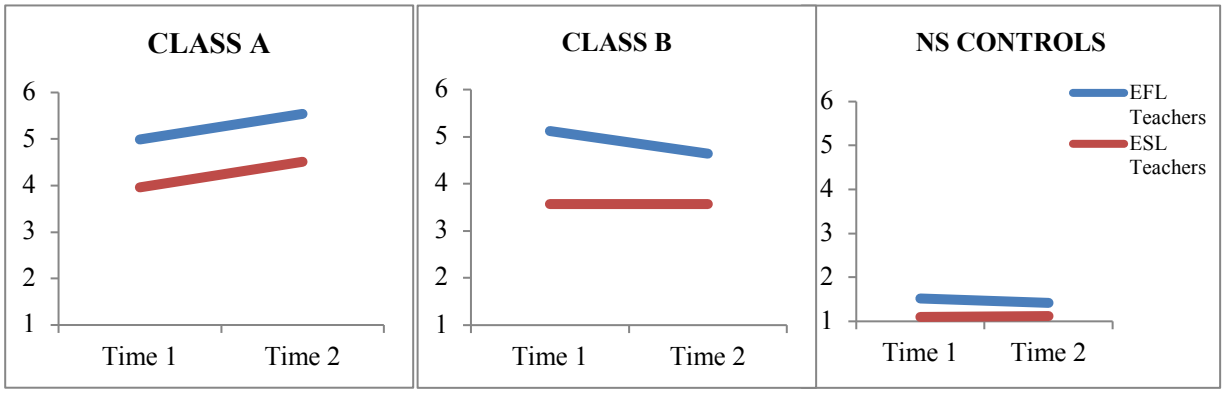


Figure 2. Means of Ratings by Rater Group over Time in Three Speaker Groups.

Qualitative

The qualitative results demonstrated commonalities between both groups of teachers' ratings, but they also showed important differences. One issue that both groups agreed that affected comprehensibility was problems with grammar and vocabulary. Mixing up tenses (e.g., not conjugating verbs in past tense) or using the wrong name for nouns (e.g., *bee* instead of *fly*) were pointed out as issues that affected comprehensibility.

The two groups of raters also mentioned that fluency and suprasegmentals affected the overall comprehensibility of the message. For example, very long pauses, or taking too much time to retrieve lexical items made L2 speech sound unnatural. Additionally, problems with word and sentence stress also made the speech more difficult to understand for both groups (see sample

comments *a* and *b* in Appendix 2). Whereas problems with fluency and suprasegmentals affected the overall comprehension of messages, problems at the segmental level hindered the comprehension of specific words. For example, substitution of one sound for another was the main issue affecting the comprehension of words (e.g., *mouth* sounded like *mouse*, or *playing* sounded like *praying*), but difficulties with the pronunciation of vowels, epenthesis, and clusters were also mentioned as factors that affected comprehensibility (see Appendix 2 *c* and *d*).

Although both groups had differences in familiarity with L1 backgrounds, many teachers recognized that such familiarity (or lack of it) may have affected their ratings. For example, some ESL teachers mentioned being stricter when rating speech that had an accent with which they were more familiar. In contrast, the EFL teachers were mostly familiar with Spanish-accented English, and they expressed that they had more difficulty understanding accents other than Spanish (see comments *e* and *f* in Appendix 2). Finally, and closely related to this last point, is the fact that the EFL teachers reported that accentedness resulted in lowered comprehensibility ratings for them; that is, they felt that they had struggled to overlook it. Although the EFL teachers claimed that they understood the messages, accentedness seemed to have forced them to concentrate harder on understanding the said messages—particularly when spoken with unfamiliar L1 accents. This did not seem to be the case with the ESL teachers, who claimed that the samples were comprehensible in spite of having different degrees of accentedness (see comments *g* and *h* in Appendix 2).

DISCUSSION AND CONCLUSIONS

Going back to the original research questions, the results demonstrated important differences and similarities between both groups of teachers: (1) *Do ESL and EFL teachers rate L2 speech comprehensibility in the same way?* According to the quantitative results, it appears that both groups followed similar tendencies in ratings, as shown by the very high levels of consistency between the two groups and by the absence of interaction between *Rater Group* over *Time*. However, the EFL teachers (all NNSs) were more severe when rating L2 speech, in contrast to the ESL teachers, who found the speech samples more comprehensible. Thus, the results here are consistent with previous research that has found NNSs to be more severe when rating L2 speech (Fayer & Krasinski, 1987; Kang, 2012; Rossiter, 2009). One possible reason for this is that NNSs may process L2 speech differently from NSs, and that they may also have slightly less robust representations for L2 words, which can affect how judgments about accentedness or comprehensibility are made (e.g. Levi, Winters & Pisoni, 2007).

The results also suggest that knowledge of different L2s and familiarity with accented speech may give certain listeners an advantage in terms of comprehensibility. Whereas the EFL teachers did not speak more than one L2 and had not spent a considerable amount of time in an L2 context, the ESL teachers spoke two or three L2s, had lived in L2 contexts, and were familiar with different accents because their classes routinely included a variety of learners with different L1 backgrounds. Therefore, it is possible that such familiarity gave the ESL teachers an advantage that helped them understand L2 speech more easily. As pointed out by Saito and Shintani (2016), more experience with L2 speech may give listeners a cognitive advantage to parse L2 speech more easily, which is reflected in the fact that it is easier for L2-experienced listeners to understand L2 speech as opposed to listeners who lack such familiarity with L2s (e.g., monolingual speakers). In the context of this study, the ESL teachers were NSs with a

higher level of experience listening to a variety of L2 accents, which may have made it easier for them to parse L2 speech. In contrast, the EFL teachers were NNSs and did not have much experience with different L2 accents. This differences in L2 background probably resulted in more difficulties for the latter group of teachers to parse the L2 speech, which they then perceived as less comprehensible.

As for the second research question (2) *On which linguistic aspects (i.e., phonological, lexical, syntactical) do ESL and EFL teachers base their ratings?*, the results also demonstrated important differences and similarities. For example, both groups indicated that syntactic and lexical problems affected the comprehension of L2 speech, a point which is consistent with previous studies demonstrating that comprehensibility is affected not only by phonological factors but also by lexical and syntactical issues (Crowther et al., 2016; Trofimovich & Isaacs, 2012). Additionally, the two groups of teachers reported that problems with suprasegmentals affected comprehension of the message as a whole, while problems at the segmental level affected comprehension of specific words. These insights expressed by both ESL and EFL teachers were also consistent with previous studies that attributed fluency and prosody as having major roles in comprehensibility (Crowther et al., 2016; Kang, Rubin, & Pickering, 2010; Saito, Trofimovich, & Isaacs, 2016) and that underscore the importance of segmentals in retrieving the meaning of specific words (see Derwing & Munro, 2015).

One of the key findings of this research is that even though the two groups of raters were ESL and EFL teachers, their ratings were in fact different when rating the same L2 speech samples. First, the ESL teachers had more experience with different accents and languages; this may have facilitated their parsing of L2 speech. This means that teacher education programs preparing teachers for pronunciation instruction should make an attempt not only to provide training in different theoretical and practical aspects of pedagogical phonology, but also to familiarize preservice teachers with different types of English accents—in both ESL and EFL contexts—to become aware of what constitutes intelligible and comprehensible L2 speech. This is particularly important in the current state of English in the world where NNSs of the language outnumber NSs (see Crystal, 1997; Graddol, 2006), and in which problems with comprehensibility and intelligibility may arise not only between speakers of the inner and outer circles, but more so among speakers of the outer and expanding circles (Jenkins, 2002; Levis, 2005).

Finally, and closely related to this last issue, is the fact that many teacher education programs do not necessarily provide a solid theoretical foundation for instructors to teach pronunciation (see Murphy, 2014; Thomson, 2013). Thus, the fact that there were differences in training in both groups of ESL and EFL teachers may raise the question of whether some of these teachers were treating both comprehensibility and foreign accent as an equal dimension—a question further implied by the EFL teachers' reports that accentedness was difficult to separate from their comprehensibility ratings. Such results may also question whether they are a reflection of classroom practices or not; that is, whether teachers try to unrealistically focus on eliminating their students' foreign accent instead of focusing on enhancing comprehensibility and intelligibility in their speech. Therefore, teacher education programs could help future preservice teachers not only by providing training on important aspects of phonetics and phonology, but also by exposing teachers to different theories of L2 phonology and L2 acquisition in general and to the appropriate practical application in class of different aspects of those theories in the teaching of L2 pronunciation.

ABOUT THE AUTHOR

Joshua Gordon is Assistant Professor in the Department of Languages and Literatures at the University of Northern Iowa. He has taught ESL, EFL, and Spanish as a foreign language. He has trained pre-service and in-service ESL and EFL teachers in the United States and in Costa Rica. He obtained a PhD in Second Language Studies from Indiana University, Bloomington. His research interests include L2 phonology and L2 pedagogy, particularly the interface between both areas in the teaching and learning of oral skills.

Contact Information

Department of Languages and Literatures
University of Northern Iowa
Bartlett Hall 2039
Cedar Falls, IA 50614
Telephone: (319) 273-2821

E-Mail: joshua.gordon@uni.edu

REFERENCES

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

Boersma, P. & Weenink, D. (2015). Praat: Doing phonetics by computer [Computer program] Version 5.3.23. Retrieved 6 May 2015 from http://www.praat.org/.

Brown, A. (1995). The effect of rater variables in the development of an occupation-specific language performance test. *Language Testing, 12,* 1-15.

Crowther, D., Trofimovich, P., & Isaacs, T. (2016). Linguistic dimensions of second language accent and comprehensibility: Nonnative listeners' perspectives. *Journal of Second Language Pronunciation, 2,* 160-182.

Crystal, D. (1997). *English as a global language*. Cambridge, UK: Cambridge University Press.

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

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

Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, *63*, 163-185.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research.* Amsterdam: John Benjamins.

Derwing, T. M., Rossiter, M. J., Munro, M. J. & Thomson, R. I. (2004). Second language fluency: Judgements on different tasks. *Language Learning, 54,* 655-679.

Fayer, J. M. & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning, 37,* 313-326.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84,* 70-79.

Gass, S. M. & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Earlbaum Associates.

Gordon Zamora, J. (2015). *Teaching and Learning L2 pronunciation: A closer look at classroom and extra-classroom factors in the development of comprehensibility in ESL learners*. Unpublished doctoral dissertation, Indiana University.

Graddol, D. (2006). *English next: Why global English may mean the end of 'English as a Foreign Language.'* London, UK: The British Council.

Harding, L. (2012). Accent, listening assessment and the potential for a shared-L1 advantage: A DIF perspective. *Language Testing, 29,* 163-180.

Hayes-Harb, R. & Hacking, J. F. (2015). Beyond rating data: What do listeners believe underlies their accentedness judgments? *Journal of Second Language Pronunciation,* 1, 1, 43-64.

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

Kang, O. (2012). Impact of rater characteristics and prosodic features of speaker accentedness on ratings of international teaching assistants' oral performance. *Language Assessment Quarterly*, *9,* 249-269.

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

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

Larson-Hall, J. (2009). *A guide to doing statistics in second language research using SPSS*. New York: Routledge.

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

MacKay, I. R., Flege, J. E., & Imai, S. (2006). Evaluating the effects of chronological age and

sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics, 27,* 157-183.

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

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

Murphy, J. M. (2014). ESL teachers are ready to teach pronunciation. In L. Grant (Ed.), *Pronunciation myths*. Ann Arbor: University of Michigan Press.

Richards, K. (2003). *Qualitative inquiry in TESOL*. New York: Palgrave Mcmillan.

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review, 65,* 395-412.

Saito, K. & Shintani, N. (2016). Do native speakers of North American and Singapore English differentially perceive comprehensibility in second language speech? *TESOL Quarterly, 50,* 421-446.

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, *37*(2), 217-240.

Simon's Cat. (2009, July 24). *Fly guy* [Video file]. Retrieved from https://www.youtube.com/watch?v=I1qHVVbYG8Y

Simon's Cat. (2010, February 10). *Snow business* [Video file]. Retrieved from https://www.youtube.com/watch?v=Tuf61OjvoPQ

Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language Learning, 41,* 177-204.

Thomson, R. I. (2013). ESL teachers' beliefs and practices in pronunciation teaching: Confidently right or confidently wrong? In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 224-233). Ames: Iowa State University.

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

Winke, P. & Gass, S. (2013). The influence of second language experience and accent familiarity on oral proficiency rating: A qualitative investigation. *TESOL Quarterly, 47,* 762-789.

APPENDIX 1

Sample Guide for Stimulated-Recall Interview (based on Hayes-Harb & Hacking, 2015).

1. *Are there particular sounds that influenced your comprehensibility rating?*
2. *Are there particular words that influenced your comprehensibility rating?*
3. *Did the rhythm and melody of the speech influence your comprehensibility rating? If so, how?*
4. *Did the intonation influence your comprehensibility rating? If so, how?*
5. *Did the fluency and speed of speech influence your comprehensibility rating? If so, how?*

APPENDIX 2

Qualitative Findings from the Stimulated-Recall Interviews.

a. The message is comprehensible in the sense that if I put all the words together I get it, but then it should be more fluent and he isn't. […] I don't think the pronunciation of specific sounds is a problem. I guess that even though he is not a native speaker, his message is clear somehow. But in my opinion pronunciation is not the problem but his pauses, his lack of fluency. –EFLS06.

b. He was very slow, not fluent at all, he keeps correcting himself to the point I barely know which word he's trying to get, the [di] (*the*) for *that*. If his fluency was ok, I don't think that would be a big deal, but he's "*so*:::" and his vowels are a little off, but yeah very unnatural pauses. –ESLS07.

c. She says that there was a "fry frying" [instead of "fly flying")], so there were problems with /r-l/ that changed the meaning, and also "mouse" [instead of "mouth"] […] some words were weird or difficult for me because we are accustomed to saying "very" [vɛɹi] and not [bɛɹi]." –EFLN05.

d. I couldn't tell what she was saying there "no eight"? and then [bʌtsə˞] [instead of "but her…"] I also had a hard time with her "thetas" [θ]. "thought" sounds like [tɔt], maybe that's what she's trying to say?? I mean her rhythm sounds natural, but her accuracy in terms of sounds just wasn't there. So sometimes instead of "cat" she's extending it to [katə], making it almost like two syllables, but I could understand it but that's just more just accent. –ESLS05.

e. However, let's say that I got the message, but again, I guess that she is not a native speaker. I'm not sure what her L1 is but she sounds like she speaks an Asian language but I don't know which one because I don't know their differences. –EFLS02.

f. I probably have some biases when I hear more of the East Asian speakers I'm probably a little harsher than I am on others who are Middle Eastern or wherever they are from. I guess his native language is Kazakh or maybe Russian. He didn't sound so Kazakh there, he sounded more Russian. –ESLS06.

g. What I feel about his accent is like he's not a native speaker, so I don't know, it is an advantage for me because I can understand, I'm accustomed to listening to this with my students every single day, but as for the message then I feel like I have to pay attention to understand. […] but it [accent] gives me the idea that I have to pay extra attention so that I don't miss anything. Because he's not going to speak like a native speaker, so probably I would have to… I don't know, like really focus on what he's saying to understand. –EFLS03.

h. He sounds fluent but he doesn't sound native-like […] there were some native speakers in the samples and their speech was like, how can I say? more fluent, yeah. Fluency is another tricky aspect to define, but yeah he doesn't sound very fluent to me.

Nagle, C. (2018). Modeling classroom language learners' comprehensibility and accentedness over time. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 17-29). Ames, IA: Iowa State University.

MODELING CLASSROOM LANGUAGE LEARNERS' COMPREHENSIBILITY AND ACCENTEDNESS OVER TIME

THE CASE OF L2 SPANISH

Charles L. Nagle, Iowa State University

Significant scholarship has focused on the development of L2 oral skills in naturalistic language learning. However, few studies have examined how instructed learners' pronunciation develops over time, despite the importance of the classroom context. This study addressed this gap by investigating L2 Spanish learners' comprehensibility and accentedness over a yearlong period. Twenty-six learners completed a sentence-building task on five occasions distributed throughout their second, third, and fourth semesters of college-level Spanish language instruction. Learners received 20 sets of images, combining the images in each set to form a simple sentence in Spanish. Eighteen native Spanish listeners rated learners' recordings for comprehensibility and accentedness using 9-point Likert scales, and mixed-effects models were fit to the ratings data using R. Learners were rated as quite comprehensible despite the presence of a moderate to strong foreign accent. Although both comprehensibility and accentedness improved over time, rates of change varied. Comprehensibility improved quickly but was subject to greater deceleration in rate of change over time. In contrast, accentedness improved steadily and did not exhibit the same degree of flattening as comprehensibility. These results intersect with work on naturalistic learners and suggest that pronunciation development may be characterized by phases of change.

INTRODUCTION

To communicate successfully in the L2, learners' pronunciation needs to be readily understandable and easy to process. Yet, despite the centrality of comprehensible pronunciation to L2 communication, there is a lack of longitudinal research on how these constructs develop over time. While it appears that the greatest changes in pronunciation are achieved during immigrants' first year of intensive L2 exposure (Derwing & Munro, 2013), little is known about classroom language learners' pronunciation development even though the language classroom is an important part of most learners' experience. According to a recent Modern Language Association report (Goldberg, Looney, & Lusin, 2015), in 2013, over 1.5 million students enrolled in postsecondary language coursework in the US alone, and over 30,000 reported a primary or secondary major in languages and literatures. Understanding how classroom learners' pronunciation changes over time stands to provide an empirical basis for curricular decisions related to whether targeted training is necessary and, if so, when it should be introduced to sustain or catalyze development. Moreover, studying classroom language learners in a foreign language context will complement previous research on immigrant populations and advanced L2

speakers in immersion contexts, resulting in a more complete understanding of how L2 speakers' pronunciation changes across a range of learning scenarios and proficiency levels.

Constructs in L2 Speech Research

L2 pronunciation scholarship has focused on global dimensions of L2 speech such as comprehensibility, fluency, and accentedness, which when taken together provide a window into how listeners evaluate L2 speakers' communicative ability. Comprehensibility represents how easy or difficult speech is to understand and can be interpreted as an index of speech intelligiblity if the latter is defined not in terms of verified transcriptions, as in previous research (e.g., Munro & Derwing, 1995), but rather as a more subjective, listener-based measure. As Saito, Trofimovich, and Isaacs (2017) point out, the operationalization of comprehensibility as intelligibility fits with real-world applications, such as language assessment scenarios where listeners are asked to provide holistic evaluations of L2 speech. In contrast to both of these constructs, accentedness is defined in reference to deviations from the pronunciation of native speech, and research has consistently demonstrated that the two are partially independent (Munro & Derwing, 1995; Trofimovich & Isaacs, 2012). Whereas accentedness judgments are predominantly tied to pronunciation-based variables, listeners seem to take into account a wider range of linguistic characteristics, including grammatical accuracy and discourse structure, when evaluating comprehensibility (O'Brien, 2014; Saito, Webb, Trofimovich, & Isaacs, 2017).

Research on Pronunciation Development

Longitudinal research on L2 pronunciation has provided evidence of nonlinear growth while highlighting distinct developmental trajectories for comprehensibility and fluency on the one hand, and accentedness on the other (Derwing & Munro, 2013; Derwing, Munro, & Thomson, 2008; Kennedy, Foote, & Dos Santos Buss, 2015). In one of the most comprehensive studies to date, Derwing and Munro (2013) examined L1 Mandarin and L1 Slavic language speakers' pronunciation learning over a 7-year period, during which participants were living and working in Canada. The Slavic language speakers' fluency and comprehensibility improved steadily over the study, but improvements to accentedness were limited to the first two years. In contrast, no gains were observed for the L1 Mandarin group, which may have been due to the fact that the Mandarin speakers demonstrated less willingness to communicate in the L2. In another study, Kennedy, Foote, and Dos Santos Buss (2015) investigated seven L2 English speakers who were attending a university where English was the language of instruction. Although participants' comprehensibility, fluency, and accentedness improved significantly over two years, individual differences in rate of change were evident. Findings from Huensch and Tracy-Ventura's (2017) study on advanced L2 Spanish learners' fluency before, during, and after study abroad reinforce these results. Twenty-seven participants completed a picture narration task on six occasions, once before departure, three times while abroad, and twice upon their return. Recordings were transcribed and analyzed for a number of speed, breakdown, and repair fluency measures. With respect to speed, results suggest that the greatest improvements were registered at the outset of study abroad, either between the presojourn recording and the first session abroad or between the latter and subsequent data points. In contrast, a more complex pattern emerged for the breakdown measures, and no significant changes were evident for the repair variables.

To summarize, even though longitudinal research on pronunciation development has grown over the past decade, most studies have targeted advanced L2 speakers, predominantly in an

immersion context. In that context, learners' pronunciation may improve rapidly at first, such that gains are concentrated at the outset of more intensive L2 contact (Derwing & Munro, 2013). However, little is known about how classroom language learners' pronunciation changes over the first few semesters of language study, despite the fact that it is precisely during this early stage when large shifts in linguistic ability might be expected. If initial trajectories set the stage for long-term attainment, it is critical that we examine classroom learners' comprehensibility and accentedness during this formative period.

METHOD

Speakers

Twenty-six L1 English-speaking students were recruited from multiple sections of a second semester college-level Spanish course at a university in the US. Although all participants were late learners placed into the same course, they varied in terms of their previous Spanish experience. The mean age of onset for the group was 14.38 (*SD* = 4.11) years, and learners reported an average of 3.35 (*SD* = 3.17) years of Spanish instruction prior to university. Participants were tested five times over a yearlong period: two sessions during their second semester, one in February and another in April; two sessions during their third semester, one in September and another in November; one session in February during their fourth semester. The language program in which participants were enrolled followed a communicative approach to language teaching. Classes met three days per week for one hour, during which students typically completed a series of scaffolded, task-based activities in pairs. Although the program focused on developing students' speaking ability, pronunciation was not part of the curriculum. Instructors reported that they did not systematically address pronunciation, but would recast students' mispronunciations—that is, provide the correct pronunciation without explicitly drawing students' attention to it—if they interfered with communication, and students likewise indicated that they did not receive any targeted pronunciation training as part of their courses.

Five participants withdrew after the first two sessions because they decided to discontinue their study of Spanish, and data from the first session was unavailable for four learners due to technical issues. Taking into account these sources of attrition, the final data set consisted of at least four observations for 21 of the original 26 participants. Randomly missing data is not problematic for mixed-effects models, the statistical approach employed in the present study, since model estimates are based on available data points without applying listwise deletion (i.e., deleting an entire case due to one missing data point).

Raters

Raters were 18 native Spanish speakers who were advanced L2 English speakers pursuing a graduate degree at a US university. They completed a background questionnaire eliciting information on their English ability, other foreign languages known or studied, and their familiarity with L2 Spanish. They reported a mean age of L2 English onset of 10.37 years (*SD* = 6.64) and estimated their English proficiency to be 7.93 (*SD* = .95; range = 6.75–9) on a 9-point scale ranging from *very poor* (1) to *extremely proficient* (9). On average, raters indicated interacting with L2 Spanish speakers about once per month (*never, n* = 6; *monthly, n* = 7; *daily, n* = 2; *more than daily, n* = 3) and assessed themselves as having moderate familiarity with the

characteristics of L2 Spanish (*M* = 5.63, *SD* = 2.99, range = 1–9 on a 9-point scale where 9 = *extremely familiar*). Eight raters reported linguistic training (i.e., having taken a course in linguistics) and nine had language teaching experience.

Speaking Task

Given that participants were novice Spanish speakers, a sentence-building task was selected for this study. Twenty sets of three images were created by combining images representing a subject, verb, and either a direct object or a location. Vocabulary was drawn from students' introductory textbook, and participants completed a computerized training module using SuperLab software prior to recording the sentences to ensure that they could easily recall the pictures associated with all vocabulary items. During the vocabulary familiarization module, participants saw an image with the word printed below it and heard the word spoken by a native Spanish speaker. After participants had reviewed the 27 picture-word pairs included in the study, they took a vocabulary quiz on which they matched images to target words by choosing the correct response from four options. Participants advanced only when they had achieved a perfect score. For the recordings, students saw one set of images per screen and were asked to form a simple sentence in Spanish based on the pictures (e.g., *Mario pinta la cocina*, 'Mario paints/is painting the kitchen'). Students were recorded individually in a sound-attenuated studio using a Shure SM10A head-mounted microphone connected to a laptop computer through an XRL to USB signal adapter.

Rating Task

The 20 sentences that each learner recorded at each testing session were similar in length and content given that they were elicited through sets of images as described above. Consequently, five per learner per session were randomly sampled using a randomization algorithm, and the peak intensity of each file was then normalized to 70 dB to prepare samples for rating. Thirty-five samples from seven native speakers of Spanish (i.e., five per native speaker) were also included. Audios were split into three blocks with each block containing files from all speakers at all testing sessions as well as native samples. Listeners were tested individually in a quiet space using noise-cancelling headphones. Comprehensibility was defined as ease of listening and accentedness as deviations from native speech, both on 9-point scales where 1 represented the best score and 9 the worst score. Ratings were carried out sequentially and counterbalanced across listeners to prevent ordering effects. In other words, listeners heard each file twice, once while rating each construct, with half of the raters evaluating comprehensibility first and the other half accentedness. Files were randomized within blocks such that each listener heard the clips in a unique order. The entire rating session lasted about one hour, and listeners were offered breaks in between blocks to prevent fatigue.

RESULTS

In this study, mixed-effects models (for an overview, Cunnings & Finlayson, 2015; Linck & Cunnings, 2015) were fit to the data using the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) of R (R Core Team, 2017). Model building followed a stepwise approach according to which fixed and random effects (by-speaker and by-rater) were progressively integrated into models and nested models were compared against one another by performing a Chi-squared test on their deviance statistics. The linear session predictor was centered on the

sample mean to reduce spurious slope-intercept correlations (Baayen, 2008), and a quadratic session predictor was computed by squaring the linear term to investigate curvature over time.

Interrater Reliability and Rater Effects

Ratings assigned to the native samples were first inspected to examine whether listeners reliably identified native speakers. One listener was removed from the data set because she assigned multiple native speakers a rating of 9, indicating that she had misinterpreted the directionality of the scale. No other anomalous ratings of the native data were observed. A continuous outcome variable was computed by aggregating ratings for the five sentences while maintaining the speaker, rater, and session factors. Interrater reliability was assessed separately for comprehensibility and accentedness using two two-way, agreement, average-measure intraclass correlation coefficients (ICC). For both constructs, the ICC was .86 (95% CI, .76 to .92). Mixed-effects models were fit to the ratings data integrating the following rater characteristics: English proficiency, percent daily Spanish use, frequency of interactions in Spanish with L2 Spanish speakers, familiarity with the characteristics of L2 Spanish, linguistics training, and teaching experience. No significant relationships emerged, indicating that raters assessed the clips similarly irrespective of their background and frequency of interaction with L2 Spanish speakers.

Developmental Models

Table 1 summarizes the taxonomy of developmental models fit. In these models, the linear session term examined linear growth over time, and the quadratic term examined curvature or quadratic growth over time. Both predictors were grand mean centered. Type refers to rating type, and interaction terms investigated whether comprehensibility and accentedness exhibited distinct rates of change over time. Random effects were included to account for by-speaker variability in development over time (e.g., adjusting the group-level coefficients slightly for each speaker), and by-rater variability in the evaluation of the speech samples.

An intercepts model was first fit including by-speaker and by-rater random intercepts (model 1) followed by two unconditional linear growth models, the first containing linear session as a fixed effect (model 2) and the second integrating by-speaker random slopes for the linear session predictor (model 3). Including quadratic session as a fixed effect significantly improved model fit (model 4), but the corresponding by-speaker random effect did not (model 5). The final unconditional growth model quantifying the effect of time on the ratings included fixed effects for linear and quadratic session, both of which were grand mean centered, by-speaker random slopes for linear session, and by-speaker and by-rater random intercepts.

Table 1

Taxonomy of Models Fit.

	Model description					Test against prior model	
	Fixed effects	Random by-subject	Random by-rater	AIC	Δ AIC	Statistic	*p*
1	intercept	intercept	intercept	15992			
2	linear			15892	–100	$\chi^2(1) = 102.38$	< .001
3	linear	linear		15870	–22	$\chi^2(2) = 24.48$	< .001
4	linear + quad.	linear		15848	–23	$\chi^2(1) = 24.72$	< .001
5	linear + quad.	linear + quad.		15846	–2	$\chi^2(3) = 7.48$	.06
6	model 4 + rating type (type)	linear		12800	–3049	$\chi^2(1) = 3049.28$	< .001
7	model 4 + type	linear + type	type	9796	–3604	$\chi^2(5) = 3014.93$	< .001
8	model 4 + linear × type	linear + type	type	9768	–28	$\chi^2(1) = 26.62$	< .001
9	model 4 + linear × type	linear × type	type	9673	–95	$\chi^2(4) = 103.18$	< .001
10	model 4 + linear × type + quad. × type	linear × type	type	9658	–15	$\chi^2(1) = 17.35$	< .001

Notes. Fixed and random intercepts were included in all models. *Linear* and *quad. (quadratic)* refer to the treatment of session as a linear or quadratic predictor of ratings. The *linear* term tested linear growth over time, and the *quadratic* term examined curvature or quadratic growth over time. Both predictors were grand mean centered. *Type* refers to rating type. The interaction terms investigated whether comprehensibility and accentedness exhibited distinct rates of change (*linear* × *type*) and curvature (*quad.* × *type)* over time.

To examine whether comprehensibility and accentedness displayed distinct developmental trajectories, the effect of rating type (nominal, two levels; comprehensibility was set as the baseline against which accentedness was compared) was modeled following a similar stepwise procedure. The first conditional model included rating type as a fixed effect (model 6) and then as a by-speaker and by-rater random effect (model 7). A linear session × type interaction was subsequently integrated as a fixed effect (model 8) and as a by-speaker random effect (model 9). The final model (model 10, summarized in Table 2) incorporated the quadratic session × type interaction as a fixed effect.

Because comprehensibility was set as the baseline level of the rating type factor, the intercept and linear and quadratic session parameters refer to comprehensibility, whereas coefficients for model parameters involving rating type represent intercepts and trajectories for accentedness when combined with the corresponding (comprehensibility) baseline term. According to model estimates, at the outset of the study, participants began with average comprehensibility and accentedness ratings (i.e., intercepts) of 2.97 and 5.82 (intercept + rating type: 2.99 + 2.83 = 5.82). The negative coefficient for the linear session term (estimate = –1.14, *SE* = .10, *t* = –11.91) indicates that participants became more comprehensible and less accented over time. However, the positive coefficient for the linear slope by rating type interaction term (estimate = .54, *SE* = .11, *t* = 4.77) suggests that accentedness improved more gradually, at a rate of approximately –.60 units per session (linear session + linear session × rating type: –1.14 + .54 = –.60). The positive coefficient for the quadratic slope term (estimate = .13, *SE* = .01, *t* = 11.20) demonstrates that rate of change decreased over time. In other words, because negative coefficients index improvement given the lower-is-better operationalization of the rating scales in this study, the positive coefficient for the quadratic session predictor acts to constrain the magnitude of the negative growth trajectory observed for the linear term. The fact that the quadratic session by rating type interaction was also negative (estimate = –.07, *SE* = .02, *t* = –4.18) shows that the flattening out of the developmental trajectory observed for comprehensibility was not as pronounced for accentedness. Finally, the statistically significant by-speaker random effect for the linear session × type term suggests that participants displayed unique trajectories in each area over time. Figure 1 displays group trajectories as thick lines and model-estimated individual trajectories as thin lines. As is evident, comprehensibility displayed a steeper initial slope that began to level off significantly by the third session. In contrast, accentedness improved at a slower rate but exhibited less curvature over time. Figure 2 plots individual trajectories in more detail to demonstrate that whereas for some individuals comprehensibility and accentedness trajectories ran parallel, suggesting similar rates of change in both areas (e.g., 4, 22, 23), for other learners, comprehensibility either outpaced accentedness or accentedness did not change much at all (e.g., 9 and 24).

Table 2

Final Model of L2 Ratings Data.

	Fixed effects			*Random effects*	
				By Speaker	By Rater
Parameters	Estimate	*SE*	*t*	*SD*	*SD*
Intercept	2.99	.34	8.72	.57	1.33
Linear session	–1.14	.10	–11.91	.31	
Quadratic session	.13	.01	11.20		
(Rating) Type	2.83	.47	5.96	.49	1.91
Linear × Type	.54	.11	4.77	.23	
Quadratic × Type	–.07	.02	–4.18		

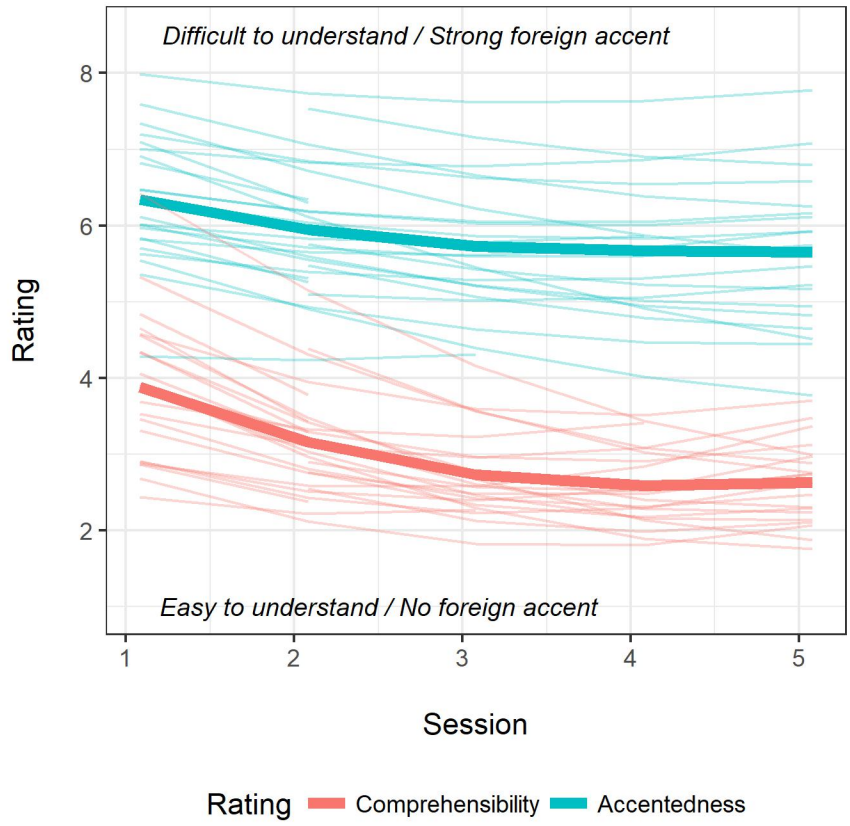


Figure 1. Development of L2 Comprehensibility and Accentedness over Five Sessions. Bold lines refer to the group trajectory and thin lines to individual trajectories.

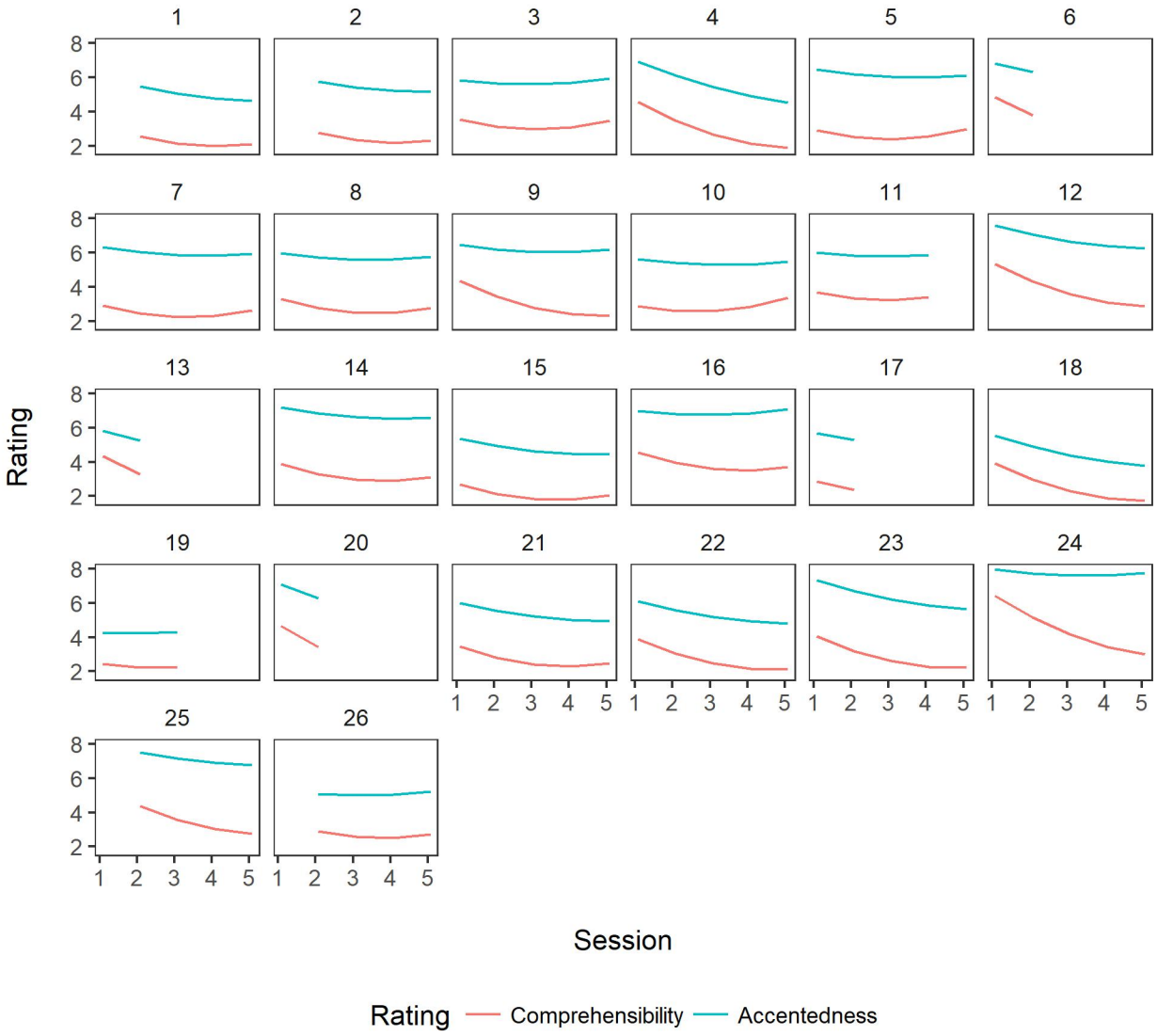


Figure 2. Individual Trajectories for L2 Comprehensibility and Accentedness over Five Sessions.

DISCUSSION

This study investigated novice L2 Spanish learners' comprehensibility and accentedness over a yearlong period encompassing data points distributed throughout their second, third, and fourth semesters of college-level Spanish language instruction. Mixed-effects models including linear and quadratic slopes were fit to examine rate of change and curvature over time in each dimension of L2 speech. At the outset of the study, learners were rated as comprehensible despite the presence of a moderate to strong foreign accent. Both comprehensibility and accentedness improved significantly while exhibiting differential rates of change. Comprehensibility improved more rapidly but was also subject to greater flattening over time, such that by the third data point the slope approached zero. On the other hand, even though

accentedness did not improve as rapidly over the first two sessions, it did not experience the same degree of leveling.

From a theoretical perspective, these findings intersect with Derwing and Munro's (2013) study on L2 immigrants living in Canada. In both cases, L2 speakers' pronunciation improved significantly with greater gains observed at the outset of the study, reinforcing the notion of a window of maximal opportunity for L2 pronunciation learning (Derwing & Munro, 2015). For classroom learners, the window of opportunity seems to extend over the first few semesters of language coursework, after which additional, targeted opportunities might be necessary to sustain development, particularly as far as accentedness is concerned. Results also intersect with Huensch and Tracy-Ventura's (2017) study tracking the development of various aspects of advanced L2 Spanish speakers' fluency while abroad. If we construct a continuum out of the speakers and observation periods contained within these three studies, then the tentative picture that emerges is one in which pronunciation development is characterized by short bursts of rapid change followed by enduring stretches of stabilization. In particular, precipitous shifts seem to occur at the onset of more intensive L2 contact, either via intensive communicative language training at university, as a result of an academic semester or year abroad, or due to relocation to an L2 environment. That is not to say that the speakers and contexts should be conflated, but rather that all three studies suggest similar developmental trajectories that may be additionally catalyzed or constrained by individual differences. However, more work is needed before a more definitive conclusion can be reached, particularly in studies tracking learners over even longer periods that encompass multiple watershed moments.

From a pedagogical perspective, results indicate that communicative language instruction can facilitate the development of more comprehensible and less accented speech in the language classroom even in the absence of pronunciation instruction. Given that accentedness improved at a slower rate, it could be beneficial to provide learners with additional targeted training to reduce accentedness. At the same time, significant individual differences in intercepts and rates of change were evident despite the fact that the 26 participants included in this study were all late learners of Spanish who had been placed into the same language course. Because the range of observed trajectories cannot be attributed to learners' previous experience or instructional factors, both of which were relatively uniform, future research should concentrate on investigating the role cognitive and attitudinal individual differences may play in the classroom context.

ABOUT THE AUTHOR

Charles Nagle (cnagle@iastate.edu) is an Assistant Professor of Spanish in the Department of World Languages and Cultures at Iowa State University. His research focuses on modeling the factors that underlie pronunciation development in L2 Spanish. His research interests include the perception-production link, individual differences in aptitude and motivation and their relationship to pronunciation learning, and the application of mixed-effects models to large data sets.

Iowa State University
World Languages and Cultures
3102 Pearson Hall
505 Morrill Road

Ames, Iowa 50011, USA
Phone: 515 294-6089

REFERENCES

Baayen, H. R. (2008). *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. New York: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). _lme4: Linear mixed-effects models using Eigen and S4_. R package version 1.1.-7. Retrieved from http://CRAN.R-project.org/package=lme4

Cunnings, I., & Finlayson, I. (2015). Mixed effects modeling and longitudinal data analysis. In L. Plonsky (Ed.), *Advancing Quantitative Methods in Second Language Research* (pp. 159-181). New York: Routledge.

Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning, 63*(2), 163-185. doi:10.1111/lang.12000

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation Fundamentals: Evidence-based perspectives for L2 teaching and research*. Philadelphia, PA: John Benjamins.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics, 29*(3), 359-380. doi:10.1093/applin/amm041

Goldberg, D., Looney, D., & Lusin, N. (2015). Enrollments in languages other than English in United States institutions of higher education, Fall 2013. Retrieved from https://apps.mla.org/pdf/2013_enrollment_survey.pdf

Huensch, A., & Tracy-Ventura, N. (2017). L2 utterance fluency development before, during, and after residence abroad: A multidimensional investigation. *The Modern Language Journal, 101*(2), 275-293. doi:10.1111/modl.12395

Kennedy, S., Foote, J. A., & Dos Santos Buss, L. K. (2015). Second language speakers at university: Longitudinal development and rater behaviour. *TESOL Quarterly, 49*(1), 199-209. doi:10.1002/tesq.212

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning, 65*(S1), 185-207. doi:10.1111/lang.12117

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning, 45*(1), 73-97. doi:10.1111/j.1467-1770.1995.tb00963.x

O'Brien, M. G. (2014). L2 learners' assessments of accentedness, fluency, and comprehensibility of native and nonnative German speech. *Language Learning, 64*(4), 715-748. doi:10.1111/lang.12082

R Core Team. (2017). R: A language and environment for statistical computing (Version 3.4.2). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, *38*(4), 439-462. doi:10.1093/applin/amv047

Saito, K., Webb, S., Trofimovich, P., & Isaacs, T. (2017). Lexical profiles of comprehensible second language speech: The role of appropriateness, fluency, variation, sophistication, abstractness, and sense relations. *Studies in Second Language Acquisition, 38*(4), 677-701. doi:10.1017/S0272263115000297

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition, 15*(4), 905-916. doi:10.1017/S1366728912000168

Sugimoto, J. & Uchida, Y. (2018). Accentedness and acceptability ratings of Japanese English teachers' pronunciation. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 30-40). Ames, IA: Iowa State University.

ACCENTEDNESS AND ACCEPTABILITY RATINGS OF JAPANESE ENGLISH TEACHERS' PRONUNCIATION

Junko Sugimoto, University of the Sacred Heart, Tokyo

Yoko Uchida, Tokyo University of Marine Science and Technology

Non-native English teachers in the EFL setting wish to serve as models for their students; however, for both teachers and students, it is not clear what "acceptable pronunciation" of English teachers is. To this end, an experiment was conducted to investigate the acceptability of non-native English teachers' pronunciation in relation to accentedness. Ten native English-speaking teachers, ten Japanese teachers, and ten Japanese students listened to an English passage read aloud by 20 Japanese speakers. The listeners rated each speaker's pronunciation in terms of "accentedness" and "acceptability as an English teacher" on a nine-point scale. A strong positive correlation was found between "accentedness" and "acceptability" within each listener group, suggesting the possibility that accentedness plays a role in judging the acceptability of non-native English teachers' pronunciation. Native English-speaking teachers tended to give higher acceptability ratings than Japanese teachers and students. There were differences among speakers for both accentedness and acceptability, although the rank order revealed a similarity for lower-ranked speakers among all three listener groups. To fully assess acceptability, other constructs such as intelligibility and comprehensibility, as well as phonetic features such as fluency, clarity, and speech rate, should be taken into account.

INTRODUCTION

Pronunciation is an essential component of language, and second/foreign language learners wish to acquire pronunciation that enables them to successfully communicate in their target language. To help learners achieve this goal, teachers play an important role. What, then, is the pronunciation goal for English teachers, especially non-native teachers in EFL settings? In pursuit of this question, this paper investigates the relationship between "accentedness" and a new construct, "acceptability." A concept similar to "acceptability," has been investigated under the term "suitability" by Boyd (2003) and Moran (2016).[1] Moran (2016) defined *perceived teaching suitability* as follows:

> an immediate impressionistic judgment on the part of the listener as to whether the speaker would be a great or poor teacher, whether he or she would help students learn, and whether the listener would like to have him or her for a teacher (p. 10).

In her study of foreign-born teachers in Sweden, Boyd (2003) found that "teacher suitability" is closely associated with language proficiency, and in particular, accentedness. The study focused on the situation where non-native speakers teach native speakers, and thus, the findings may not be directly comparable to the English teaching context in Japan, where non-native speakers teach non-native speakers who share the same mother tongue. Nevertheless,

[1] We are grateful to Alyssa Kermad and Meghan Moran for their input on relevant resources on teacher acceptability/suitability.

Boyd's findings are insightful in that accent is found to play an important role in evaluating teacher suitability.

The relationship between accentedness and teacher suitability was also suggested in the authors' studies. To investigate Japanese teachers' ideas on their own pronunciation goals, Uchida and Sugimoto (2016, 2017) conducted studies using questionnaires targeting both in-service (i.e., currently active) and preservice (i.e., under training) teachers and found more than 80% believed it desirable for teachers to acquire native-like pronunciation. A follow-up qualitative survey on preservice teachers revealed that this belief stemmed from the perceptions of teachers' roles as an important source of input and a good model for students. Many described teachers' pronunciation should be "higher level/better (than students)" or "correct" (Uchida & Sugimoto, 2017).

Although many researchers now agree that pronunciation goals for non-native speakers has shifted from native-like pronunciation to intelligible pronunciation (Celce-Murcia, Brinton, & Goodwin, 2010; Derwing & Munro, 2005; Levis, 2005), this idea is not prevalent among non-native teachers worldwide (Jenkins, 2007), including Japan. One possible reason is that the level of pronunciation non-native English teachers should aim at is not clear, and this has resulted in a deep-rooted inclination toward acquisition of native-like pronunciation.

Thus, we decided to further investigate the relationship between the two constructs: "accentedness" and "acceptability." The definition of "accentedness" in our study follows that of Derwing and Munro (2015, p. 175): "the extent of difference perceived by speakers of one linguistic variety when listening to speakers of other varieties." The construct of "acceptability" in this study refers to "how acceptable a non-native speaker's pronunciation is as a teacher." By asking three listener groups (native English-speaking teachers, Japanese teachers, and Japanese students) to participate in rating tasks, the following two research questions were explored: *(i) Do "accentedness" and "acceptability" ratings differ from each other within a listener group?* and *(ii) Do "accentedness" and "acceptability" ratings differ across the three listener groups?* We hypothesized that listeners' evaluations of "acceptability" are influenced by "accentedness," and that while listeners with different background may evaluate "accentedness" in the same way (cf. Munro et al., 2006), they may evaluate "acceptability" differently.

METHOD

Materials

In a typical English class in a Japanese junior high school, a teacher is often required to read aloud a passage from a textbook to provide a model pronunciation. Having this teacher's role in mind, materials for the experiment were selected from an English textbook (Togo et al., 2013, p. 80). The passage is an excerpt from the speech originally delivered by Severn Suzuki at the UN Earth Summit in 1992.

- *I am here to speak for starving children around the world.*
- *I am afraid to breathe the air because I don't know what chemicals are in it.*
- *Did you have to worry about these things when you were my age?*

These three sentences were specifically chosen because they include segments that are often said to be difficult for Japanese speakers (e.g., /ɚː, ɑɚ, l, r, θ, ð/) and have sentence structures that can be used to test appropriate rhythm and intonation.

Speakers

For the current experiment, we prepared 20 speech samples, 13 of which were recordings made by preservice teachers enrolled in a phonetics course for a teacher's certificate. The rest were supplemented from a pool of recordings by 57 students who recorded the same passage as an assignment for another class. Recordings with hesitations or misread words, which are reported to have a great impact on listeners' evaluations (Matsuura, Chiba, & Ara, 2012), were not chosen because those features were not our main focus. All the students were from two universities in Tokyo, and their profiles are as follows: gender (11 female, 9 male); CEFR level measured with the Cambridge English Placement Test (A2 = 6, B1 = 7, B2 = 6, C1 = 1). The script was given in advance for speakers to understand its meaning and to practice reading.

Listeners

Three groups, each consisting of ten listeners, were recruited to participate in the experiment: ten native English-speaking teachers (ET), ten Japanese teachers (JT), and ten Japanese students (JS). Of the ten native English-speaking teachers (age range: 37–56, *M* = 44.7; 3 female, 7 male), four were from the United States, two from Australia, and one each from the United Kingdom, Canada, Ireland, and Spain.[2] All ten teachers, having had experience of both living and teaching in Japan, were familiar with the Japanese accent; their mean length of stay in Japan was 13.4 years (range: 4.8–20 years), and the mean length of teaching experience in Japan was 12.4 years (range: 4.8–20 years). Ten Japanese teachers (age range: 34–60, *M* = 47.9; 9 female, 1 male) were all active teachers at the college or university level in Japan, but none specialized in phonetics/phonology or pronunciation teaching. Last, the ten Japanese students (age range: 20–21, *M* = 20.4; 10 female) were all university students majoring in English with the following level of English proficiency: CEFR A2 = 7, B1 = 1, and B2 = 2.

Procedure

Before starting the experiment, all listeners read the test passage. The Japanese listeners (JT and JS) were also introduced to the model recording from the CD that accompanied the original textbook.[3] After that, the listeners were provided with the instructions.

In the experimental session, the listeners were first asked to listen to the 20 recordings of Japanese speakers in a randomized order and evaluate the accentedness of each speech sample on a nine-point scale (1 = *heavily accented*, 9 = *not accented*). Next, they listened to the same 20 speakers again in a randomized order, and this time they were asked to evaluate how acceptable the speaker's pronunciation was as an English teacher, again on a nine-point scale (1 = *not acceptable*, 9 = *very acceptable*). The listeners were allowed to listen to each speech sample only once. To familiarize listeners with the scaling, a practice session with three trials preceded each rating task.

After completing the experiment, all listeners filled out a language background questionnaire. In addition, they answered two questions on the rating tasks: "On what basis did you evaluate 'accentedness' of the speaker's pronunciation?" and "On what basis did you evaluate

[2] One participant was born in Spain but we confirmed that his first language was English.
[3] We decided to provide Japanese listeners with a model pronunciation simply because some of them (mostly students) might not have known the correct pronunciation of some words (e.g., *breathe*).

'acceptability' of the speaker's pronunciation as that of an English teacher?"

Each listener from the ET and JT groups received an experiment package (a consent form, instruction sheets, and a USB stick with audio files) by mail and was asked to send it back upon completion. In the case of the JS group, each listener individually took the experiment in a quiet room in the presence of the first author. All listeners were offered payment. The experiment took less than 30 minutes, and the listeners were advised to take a few minutes' break after completing the accentedness rating.

RESULTS

The final data set did not include ratings for two speakers; one listener from the ET group left one acceptability rating blank, and another chose an inappropriate value for one accentedness rating. These two missing values were replaced by the median of the other speakers' ratings given by the listener in question.

Correlations between Accentedness and Acceptability within Each Listener Group

After confirming high inter-rater reliability (Cronbach's $\alpha > .80$) between the listeners within each group, mean scores of accentedness and acceptability ratings were obtained from the three groups (Table 1). A strong positive correlation was confirmed for all three groups between accentedness and acceptability ratings: *Pearson* $r(18) = .93$ for NT, $r(18) = .87$ for JT, and $r(18) = .93$ for JS, $p < .001$ (Figure 1). Notably high correlations for all three groups indicate that the listeners' ratings for accentedness and acceptability were fairly similar.

Table 1

Means and Standard Deviations of Accentedness and Acceptability Ratings by the Three Listener Groups

	ET	JT	JS
Accentedness	5.28 (1.51)	5.48 (1.66)	5.46 (1.65)
Acceptability	6.21 (1.64)	5.31 (2.35)	5.32 (2.37)

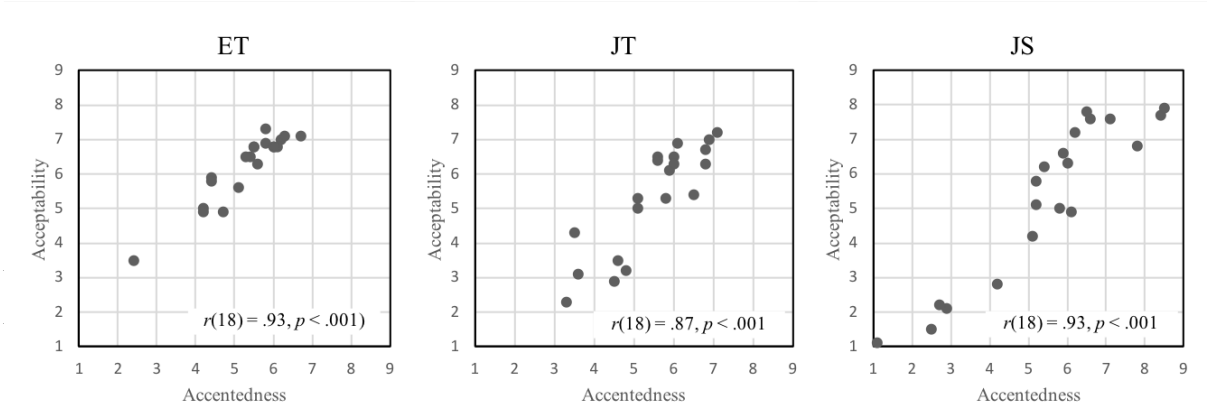


Figure 1. Correlations between accentedness and acceptability by the three listener groups.

Accentedness and Acceptability Ratings across Listener Groups

A two-way ANOVA was conducted to evaluate the effects of speaker and listener group. The independent variables were speaker, listener group, and the interaction between speaker and listener group, and the dependent variables were accentedness and acceptability ratings. As for accentedness, the main effect for speaker yielded $F(19, 513) = 24.40$, $p < .05$, indicating a significant difference between 20 speakers. The rating was the highest for S02 ($M = 7.2$, $SD = 1.58$) (Table 2). In contrast, the main effect of listener group was non-significant, $F(2, 27) = .14$, $p > .05$. The interaction effect was significant, $F(38, 513) = 2.70$, $p < .05$.

Table 2

Accentedness Ratings by the Three Listener Groups

Code	CEFR	*M (SD)*			
		ET	JT	JS	Overall
S01	B1	6.7 (1.57)	6.5 (1.72)	5.9 (2.42)	6.4 (1.90)
S02	B2	6.3 (1.34)	6.8 (1.75)	8.4 (0.70)	7.2 (1.58)
S03	C1	6.2 (1.69)	6.1 (1.45)	6.2 (1.87)	6.2 (1.62)
S04	B2	6.1 (1.52)	5.9 (1.10)	6.0 (1.76)	6.0 (1.44)
S05	B2	6.0 (1.83)	6.8 (2.30)	6.5 (2.64)	6.4 (2.22)
S06	B1	6.0 (1.41)	5.6 (1.65)	6.1 (1.66)	5.9 (1.54)
S07	B1	5.8 (1.81)	6.0 (1.89)	7.8 (0.91)	6.5 (1.80)
S08	B1	5.8 (1.14)	6.9 (1.85)	6.6 (2.22)	6.4 (1.79)
S09	B2	5.6 (1.58)	7.1 (1.66)	8.5 (0.71)	7.1 (1.80)
S10	A2	5.5 (1.90)	5.1 (1.97)	5.1 (2.42)	5.2 (2.05)
S11	A2	5.5 (1.43)	4.6 (2.01)	5.2 (2.04)	5.1 (1.83)
S12	B1	5.4 (1.84)	5.6 (1.58)	5.2 (1.62)	5.4 (1.63)
S13	B2	5.4 (1.35)	5.8 (1.40)	5.8 (1.48)	5.7 (1.37)
S14	B2	5.1 (1.37)	6.0 (0.94)	7.1 (0.88)	6.1 (1.34)
S15	B1	4.7 (1.42)	5.1 (1.37)	5.4 (1.78)	5.1 (1.51)
S16	A2	4.4 (1.58)	4.8 (1.81)	2.7 (2.06)	4.0 (1.99)
S17	B1	4.4 (1.08)	3.5 (1.35)	4.2 (2.15)	4.0 (1.59)
S18	A2	4.2 (1.62)	3.6 (1.58)	2.5 (1.18)	3.4 (1.59)
S19	A2	4.2 (1.62)	4.5 (2.32)	2.9 (2.13)	3.9 (2.10)
S20	A2	2.4 (1.08)	3.3 (1.57)	1.1 (0.32)	2.3 (1.41)

Note: The speaker codes (S01 to S20) are assigned based on the ET group's rank order of accentedness.

With regard to acceptability, the main effect for speaker yielded $F(19, 513) = 30.71$, $p < .05$, indicating a significant difference between 20 speakers. The rating was the highest for S05 ($M = 7.1$, $SD = 1.95$), S08, ($M = 7.1$, $SD = 1.32$) and S09 ($M = 7.1$, $SD = 1.94$) (Table 3). The main effect for listener group yielded $F(2, 27) = 3.96$, $p < .05$, indicating a significant difference between listener groups. The ET group showed the highest acceptability rating of all three groups ($M = 6.21$, $SD = 1.64$) (Table 1). The interaction effect was significant, $F(38, 513) = 4.26$, $p < .05$.

Table 3

Acceptability Ratings by the Three Listener Groups

Code	CEFR	*M (SD)*			
		ET	JT	JS	Overall
S01	B1	7.1 (1.29)	5.4 (1.51)	6.6 (2.01)	6.4 (1.73)
S02	B2	7.1 (1.52)	6.3 (2.98)	7.7 (1.83)	7.0 (2.21)
S03	C1	7.0 (1.16)	6.9 (1.10)	7.2 (0.79)	7.0 (1.00)
S04	B2	6.8 (1.14)	6.1 (1.52)	6.3 (1.89)	6.4 (1.52)
S05	B2	6.8 (1.69)	6.7 (2.26)	7.8 (1.87)	7.1 (1.95)
S06	B1	6.8 (1.23)	6.4 (1.17)	4.9 (1.79)	6.0 (1.61)
S07	B1	7.3 (0.82)	6.5 (2.01)	6.8 (1.69)	6.9 (1.57)
S08	B1	6.7 (1.64)	7.0 (1.15)	7.6 (1.07)	7.1 (1.32)
S09	B2	6.3 (1.49)	7.2 (2.44)	7.9 (1.60)	7.1 (1.94)
S10	A2	6.8 (1.32)	5.0 (1.89)	4.2 (2.44)	5.3 (2.17)
S11	A2	6.8 (0.63)	3.5 (1.78)	5.1 (1.52)	5.1 (1.93)
S12	B1	6.5 (1.27)	6.5 (1.51)	5.8 (2.30)	6.3 (1.72)
S13	B2	6.5 (1.08)	5.3 (1.83)	5.0 (2.11)	5.6 (1.79)
S14	B2	5.6 (1.35)	6.3 (1.34)	7.6 (0.84)	6.5 (1.43)
S15	B1	4.9 (1.79)	5.3 (2.11)	6.2 (1.81)	5.5 (1.93)
S16	A2	5.8 (1.14)	3.2 (1.81)	2.2 (1.55)	3.7 (2.13)
S17	B1	5.9 (1.52)	4.3 (2.45)	2.8 (1.23)	4.3 (2.17)
S18	A2	5.0 (1.70)	3.1 (1.97)	1.5 (0.97)	3.2 (2.12)
S19	A2	4.9 (1.60)	2.9 (1.91)	2.1 (1.85)	3.3 (2.10)
S20	A2	3.5 (1.84)	2.3 (1.64)	1.1 (0.32)	2.3 (1.70)

For both accentedness and acceptability, the interaction effect was significant. To further investigate differences between speakers, accentedness ratings by the three groups were sorted in terms of ranking. With the exception of S02 (ranked the second highest by both ET and JS), variability in accentedness was observed among the highly ranked speakers. One

example is S09, who was ranked the highest by both JT and JS groups but was ranked only ninth by the ET group. In contrast, the lowest rankings were more or less consistent across the listener groups. S18, S19, and S20 were all within the lowest four for all three groups.

Turning to acceptability, inconsistency was also found regarding the speakers who received high rankings across the listener groups. The speakers who were ranked high by ET were not always ranked high by JT and JS, and the rankings of JT and JS did not coincide either. The only exceptions were S02 (ranked the second highest by ET and the third highest by JS) and S09 (ranked the highest by both JT and JS). In contrast, the lowest rankings of acceptability were consistent. Here again, S18, S19, and S20 were among the lowest four in all three groups.

Comments on Accentedness and Acceptability

At the end of the experiment, the listeners described the basis on which they evaluated accentedness and acceptability. As expected for accentedness, many of them wrote well-known segmental and suprasegmental features that are said to be problematic for Japanese learners (e.g., /l/-/r/, rhythm, flat intonation). For acceptability, four (ET = 2, JT = 1, JS = 1) listeners wrote that they used the same criteria when judging accentedness and acceptability, six (ET = 3, JT = 2, JS = 1) wrote that they used similar criteria but incorporated other features as well, and the remaining 20 wrote that they used different criteria. Interestingly, the four listeners who mentioned speed were all Japanese students; three wrote that slower rate is preferable, whereas one mentioned that the rate should not be too slow. The features that repeatedly appeared in the comments to describe accentedness and acceptability are summarized in Figure 2.

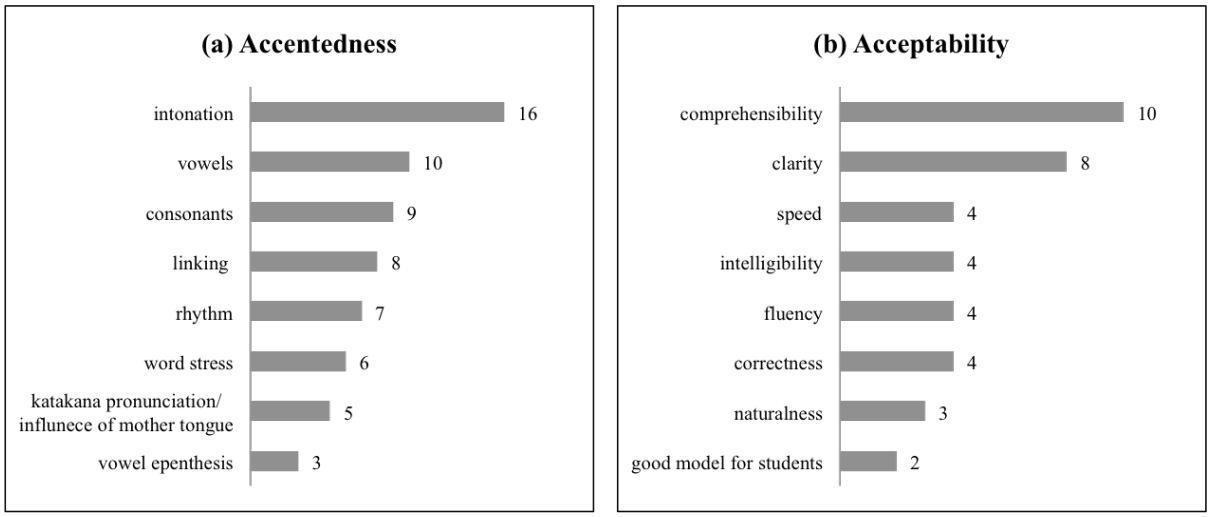


Figure 2. Comments on accentedness and acceptability. The numbers indicate the frequency. "Katakana pronunciation" in accentedness is a term often used to describe English with a typical Japanese accent. In acceptability, comments such as "easy to understand" and "can be correctly understood" were included under "comprehensibility" and "intelligibility" respectively following the definitions given by Derwing and Munro (2015).

DISCUSSION

Based on the findings, this section answers the research questions. A discussion on the concept of acceptability follows.

Research Question 1: ***Do "accentedness" and "acceptability" ratings differ from each other within a listener group?***

High correlations between accentedness and acceptability were found regardless of listener groups. In fact, one-third of listeners commented that they judged both accentedness and acceptability on the same or similar criteria, and this was reflected in the high correlations between the two constructs. Less accented speech was considered more acceptable, and more accented speech was considered less acceptable by all three groups. As anticipated, the results suggest that accentedness may be one factor that affects judgment of acceptable pronunciation as a teacher.

One possible problem with the current study, however, was the effect of task order (i.e., accentedness ratings followed by acceptability ratings). This could have biased the listeners' judgments on acceptability, which may have led to conspicuously high correlations between the two. To provide more evidence for the strong correlation between accentedness and acceptability ratings, it will be necessary to recruit the same numbers of listeners for each group and counterbalance the task order.

Research Question 2: ***Do "accentedness" and "acceptability" ratings differ across the three listener groups?***

A two-way ANOVA revealed a significant difference across listener groups for acceptability but not for accentedness. We found a higher acceptability rating by the ET group, which may suggest their more lenient attitude toward acceptable pronunciation of non-native teachers. For both accentedness and acceptability, consistency was observed in speakers who received low ratings across the three groups. This may imply that a clear image of "strongly accented pronunciation" and "unacceptable pronunciation as a teacher" is shared among all the listener groups. By conducting a preliminary observation in an attempt to investigate the phonetic characteristics that low-rated speakers share, we found that these speakers demonstrated problems in both segmentals and suprasegmentals and spoke with a slow speech rate. The three speakers (S18, S19, S20) who received low ratings had CEFR A2 level. Nevertheless, we cannot say anything conclusive about the relation between English proficiency and accentedness/acceptability, because other A2 level speakers did not receive lower ratings compared to speakers with higher CEFR level. Further analyses will be necessary to reach definitive conclusions.

In contrast, although the ratings themselves did not show much difference, variability was found for speakers who received high rankings across the three groups. To introduce one example, S09 was ranked the highest by JT and JS in terms of both accentedness and acceptability, but not by native English-speaking teachers. The notable characteristic of this speaker was her use of a wide pitch range. To the Japanese teachers and students, compared to a flat and monotone speech often criticized in Japanese speech, her excessive use of pitch range was regarded as favorable and may have sounded more English-like. However, S09's speech was found to contain segmental errors (confusion of /l/-/r/ and vowel epenthesis), and this may have resulted in the ET group's lower judgment of her speech in terms of both accentedness and acceptability. This is in line with Riney, Takagi, and Inutsuka's (2005) findings in that Japanese listeners tended to judge Japanese accent based on suprasegmentals, whereas native listeners did so on segmentals.

Defining Acceptability

Although accentedness plays a role in evaluating acceptability of non-native teachers'

pronunciation, the comments provided by the listeners in this study revealed that not only phonetic features but also other features impact acceptability. As introduced in Figure 2, listeners' evaluations of acceptability were based on various features such as comprehensibility, clarity, speed, and intelligibility.

The fact that "acceptability" may be a composite of diverse features was also suggested in a qualitative survey of Japanese preservice teachers (Uchida & Sugimoto, 2017). Some mixed views on "teacher's pronunciation" were reported. For example, some voiced that it is important for teachers to possess pronunciation that sounds "familiar" and "attainable" to students. One preservice teacher commented that pronunciation that is clear to students, even with a deliberate Japanese accent, is suitable when teaching Japanese students. The concept of "acceptability" needs to be further investigated, not only in relation to accentedness but also to other constructs such as intelligibility and comprehensibility.

Limitations

Although the current experiment was carefully designed and conducted, a few limitations in methodology should be noted. Experimental conditions were not fully controlled across listener groups; for example, only JT and JS were instructed to listen to the model recording. In addition, because of participant availability and time constraints, both ET and JT were not monitored during the experiment. Thus, there were possibilities such as listeners not fully following the instructions.

Another limitation is data interpretation. The differences among speakers were discussed based on rankings, instead of ratings, in an attempt to compare the evaluation across listener groups. Another reason for using rankings was to avoid directly comparing the absolute statistical values of ratings between groups that differ in both means and standard deviations. However, it is true that the values were similar in a few cases even though the rankings were different. To further clarify the difference between accentedness and acceptability, and to expand this experiment to more diverse listener groups (e.g., those who are not familiar with Japanese-accented English) or different constructs (e.g., intelligibility, comprehensibility), these limitations should be overcome.

CONCLUSION

This study found strong correlations between accentedness and acceptability ratings in all three listener groups: native English-speaking teachers, Japanese teachers, and Japanese students. In addition, listeners with different backgrounds were found to evaluate acceptability differently, but all listener groups seemed to have a shared image of "strongly accented pronunciation" and "pronunciation not acceptable as a teacher."

To fully evaluate teachers' pronunciation, we need to take into account speech styles (analyzing not only read speech but also spontaneous speech), phonetic features (both segmentals and suprasegmentals), and global features (e.g., clarity, fluency). Although the concept of "acceptable pronunciation as an English teacher" is not yet fully understood, setting a clear pronunciation goal for non-native teachers will be an important step toward devising a framework to assess and improve their pronunciation.

ABOUT THE AUTHORS

Junko SUGIMOTO is currently an associate professor at the University of the Sacred Heart, Tokyo. She has an MA in Phonetics from University College London. Her research interests

include acquisition of rhythm and intonation by EFL learners of English, teaching pronunciation, and intelligibility. She is currently focusing on how to effectively teach selected segmentals and suprasegmentals that affect intelligibility to Japanese learners of English.

University of the Sacred Heart, Tokyo
4-3-1 Hiroo, Shibuya-ku, Tokyo 150-8938, Japan
sugimoto@u-sacred-heart.ac.jp

Yoko UCHIDA studied at Tokyo University of Foreign Studies and obtained her PhD in Humanities in 2001. She is currently a professor at Tokyo University of Marine Science and Technology. Her research interests include but are not limited to perception of English sounds by Japanese speakers and Maritime English. One of her current topics is pursuing the intelligibility levels expected of different groups of native Japanese speakers who use English in different contexts and for different purposes.

Tokyo University of Marine Science and Technology
2-1-6 Etchujima, Koto-ku, Tokyo 135-8533, Japan
uchidayo@kaiyodai.ac.jp

REFERENCES

Boyd, S. (2003). Foreign-born teachers in the multilingual classroom in Sweden: The role of attitudes to foreign accent. *International Journal of Bilingual Education and Bilingualism, 6*(3&4), 283-295.

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

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*, 379-397. doi:10.2307/3588486

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals.* Amsterdam: John Benjamins Publishing Company. doi:10.1075/lllt.42

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

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*, 369-377. doi:10.2307/3588485

Matsuura, H., Chiba, R., & Ara, S. (2012). International intelligibility of learner English. *Shogaku Ronshu, 81*, 63-74.

Moran, M. (2016). *Arizona teachers' speech: Phonological features and listener perceptions* (Unpublished doctoral dissertation). Northern Arizona University, Flagstaff, Arizona, US.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition, 28*, 111-13. doi:10.1017/S0272263106060049

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

39, 441-466.

Togo, K. et al. (2013). *Columbus 21.* Vol. 3. Tokyo: Mitsumura Tosho.

Uchida, Y., & Sugimoto, J. (2016). A survey of Japanese English teachers' attitudes towards pronunciation teaching and knowledge on phonetics: Confidence and teaching. *Proceedings of ISAPh2016: Diversity in Applied Phonetics,* 38-42. doi:10.21437/ISAPh.2016

Uchida, Y., & Sugimoto, J. (2017, June). *Towards the implementation of ELF-oriented pronunciation teaching in Japan.* Paper presented at the meeting of ELF 10 and Changing English: 10th Anniversary Conference of English as a Lingua Franca, Helsinki, Finland.

De Weers, N., & Munro, M. (2018). The role of duration in Japanese speakers' productions of English vowels. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 41-53). Ames, IA: Iowa State University.

THE ROLE OF DURATION IN JAPANESE SPEAKERS' PRODUCTIONS OF ENGLISH VOWELS

Noortje de Weers, Simon Fraser University

Murray J. Munro, Simon Fraser University

Japanese uses length as a primary acoustic cue to distinguish some pairs of vowels, whereas North American English varieties rely mainly on spectral differences. While many studies investigating Japanese learners' acquisition of English vowels have examined acoustic features, only a handful have focused in detail on durational patterns. This study adds to the latter body of research by investigating the carry-over of Japanese durational features into English. To this end, native Japanese and Canadian English speakers' productions of CVC words were elicited using a picture-naming task. A comparison of duration data for /i/, /ɪ/, /u/, /ʊ/ in three pre-consonantal contexts (/k/, /t/, /d/) revealed that the Japanese speakers' tense vowels were significantly longer than their lax counterparts regardless of the following consonant, while the native English speakers generally did not show such a pattern. The Japanese speakers furthermore displayed native-like lengthening of vowels before a voiced stop. Nonetheless, contrary to hypothesis, an assessment by native English listeners yielded no evidence that the exaggeration of durational differences for tense/lax contrasts adversely affected vowel intelligibility.

INTRODUCTION

Various studies have examined the effect of first language (L1) on Japanese learners' acquisition of English vowels (Ingram & Park, 1997; Lee, Guion, & Harada, 2006; Oh et al., 2011; Tsukada, 1999). While acoustic features have been addressed in this work, only a handful of investigations (e.g., Liu, Jin & Chen, 2014; Tsukada, 2009) have focused specifically on durational patterns. This is surprising, considering that the two languages differ in their use of duration in a number of respects. For example, while Japanese employs duration as a primary acoustic cue to differentiate certain pairs of vowels and consonants, North American varieties of English make vowel distinctions mainly at the spectral level, with duration playing a relatively minor role in tense-lax pairs (Flege, Bohn, & Jang, 1997; Hillenbrand, Clark, & Houde, 2000). Production data show that English tense vowels are, on average, longer than lax vowels. However, perceptual research indicates that native English speakers distinguish such tense-lax pairs as /i/-/ɪ/ and /u/-/ʊ/ at high rates of accuracy when duration differences are neutralized (Hillenbrand et al., 2000). Moreover, duration differences alone do not cue perceptual distinctions. It follows that L2 users of English must learn to perceive and produce the spectral differences in these distinctions in order to successfully acquire the English vowel inventory.

In Japanese, length is phonemic, and, in contrast to English, five vowel pairs are distinguished primarily on the basis of duration, including /i:/ - /i/ and /u:/- /u/. Given the phonemic status of duration in their L1, coupled with the fact that English does, to some degree, exhibit length differences in such pairs, speakers of Japanese may be prone to misinterpret English tense-lax

distinctions as duration-based contrasts. In fact, a number of L2 speech studies have pointed to learners' difficulties with English lax vowel accuracy. In Munro (1993), for instance, Arabic speakers exaggerated tense-lax duration differences and were rated as producing lax /ɛ/ much less well than other English front vowels. Munro and Derwing (2008) observed that lax /ɪ/ was by far the least intelligible vowel produced by English speakers from Mandarin and Slavic backgrounds. And in a study of Italian immigrants in Canada, Munro, Flege and MacKay (1996) observed poor performance on the same vowel, as well as on /ʊ/. Finally, both Bohn (1995) and Wang and Munro (2004) observed that Mandarin speakers had difficulty perceiving the difference between English /i/ and /ɪ/, in part because they relied inappropriately on length differences.

A production study that investigated the transfer of Japanese duration phenomena into learners' English tense-lax vowel pairs was Tsukada (2009), who found that the Japanese speakers tended to exaggerate the duration distinction between English tense /i/ and lax /ɪ/ in comparison to Australian English speakers. In particular, her native English groups produced /ɪ/-/i/ duration ratios of 0.69 and 0.67 in a /p/ context, while the Japanese group exhibited a considerably larger duration difference with a ratio of only 0.55. Such exaggeration may reflect incomplete acquisition of the distinction, though Tsukada recommended that future work should establish the perceptual relevance of their production patterns through evaluation by native listeners.

Another pertinent difference between Japanese and English concerns the lengthening of vowels before tautosyllabic voiced obstruents. While this phenomenon is pervasive in English and serves as an important perceptual cue to obstruent voicing in syllable codas (Raphael, 1972), Japanese allows very few coda consonants, none of which are voiceless (Vance, 1987). Consequently, Japanese has no phonemic distinction between voiced and voiceless consonants in syllable-final position. On the basis of production data showing relatively small differences in vowel duration in V+t and V+d syllables, Tsukada (2009) concluded that her Japanese speakers had not acquired this duration cue.

THE PRESENT STUDY

This study will add to the body of research on English L2 vowel acquisition by further investigating the carryover of Japanese length phenomena into English. We extended Tsukada's original analysis by (1) adding the two high back vowels /u/ and /ʊ/ to determine whether her findings would be paralleled in another tense-lax distinction, (2) including one additional consonantal context (/k/), and (3) focusing on a different English variety, Canadian English. To probe the potential effect of exaggerated duration contrasts on intelligibility (as suggested by Tsukada), we also added a vowel identification assessment by native English listeners.

Expectations and Hypothesis

On the basis of the previous findings on Japanese acquisition of English vowels, we undertook this study with two expectations: (E1) that Japanese speakers would tend to exaggerate the duration differences between the English tense/lax vowel pairs /i/-/ɪ/ and /u/-/ʊ/, and (E2) that Japanese speakers would not produce a native-like duration difference in vowels preceding tautosyllabic voiced vs. voiceless consonants.

We also tested one specific new hypothesis: (H1) that exaggeration of the durational contrast in tense-lax pairs would occur at the expense of accurate vowel quality. In particular, such exaggeration might indicate a general lack of success in acquiring the vowel distinctions. Learners

might, for instance, have inadequate articulatory control of the sounds because they have focused too much on the duration differences. Furthermore, it is possible that the more a learner of English uses duration to distinguish tense and lax vowels, the less the speaker accurately produces the necessary spectral differences. Since some degree of spectral accuracy is essential for listeners to perceive tense/lax distinctions, the less intelligible these individuals' vowel productions will be.

METHODS

Stimuli and speakers

The speakers were 24 (6 male, 18 female) Japanese speakers of English who had been living in Western Canada for a mean of 2.4 years (SD = 1.97). All had come to Canada as young adults, and all had high oral English proficiency as evidenced by the fact that they were enrolled in or had recently been enrolled in English-speaking post-secondary institutions in Canada. In addition, they had all studied English in the Japanese school system, typically beginning at age 12 or 13.

Procedures

Recordings of 26 CVC words (see Table1) were made in a sound-treated room via a picture naming task. No spoken model was provided. On the basis of randomized presentations of the pictures, each speaker uttered the target items within the sentence frame "Now I say __." Three productions of each item were digitally recorded. For baseline data, the same words were also elicited, using an identical procedure, from six male and three female native Canadian English speakers who had all grown up in Western Canada.

Table 1

Rhymes and target words used to elicit the four high vowels

Rhyme	Target Words	Rhyme	Target Words
/it/	heat, seat, feet	/ut/	boot, suit
/ik/	cheek, speak	/uk/	Luke, toque[1]
/id/	read, feed	/ud/	food
/ɪt/	hit, sit	/ʊt/	put, foot
/ɪk/	chick, kick, sick	/ʊk/	book, cook, look
/ɪd/	kid, lid	/ʊd/	good, wood

[1] Pronounced /tuk/ and sometimes spelled "tuque." Widely used in Canada to refer to a close-fitting winter hat.

The four high vowels under consideration were the front vowels /i/ and /ɪ/ and the back vowels /u/ and /ʊ/, produced in the pre-consonantal voiceless contexts /k/ and /t/, and the voiced context /d/. We opted not to include final-/g/ items because of the rarity of high-frequency English words with /ig/ rhymes. In anticipation of future analyses of word-based effects on duration (not to be reported here), we included more than one word for most of the rhymes. Although we attempted to identify multiple common target words that would be familiar to the talkers and that would have similar consonants for each rhyme, in some instances this proved impractical, either because the rhyme

type was rare, or because the possible targets could not be easily depicted with pictures. In the case of /ud/, for instance, we identified only one suitable target word ('food').

Vowel duration measurements

With the help of spectral and waveform displays in *Praat* (Boersma & Weenink, 2017), the onsets and endpoints of the vowels were located and demarcated with cursors by a research assistant. The chief criteria for identifying vocalic stretches were the presence of quasi-periodicity in the signal, along with energy in the lower formants (F1 and F2). Cursor locations were visually inspected by the authors, and tokens were re-measured when inaccuracies were noted. A *Praat* script was then used to extract the vowel durations to the nearest 1 ms.

Intelligibility assessments

The intelligibility evaluations were provided by two phonetically-trained native-speaking Canadian judges who each completed a forced-choice identification task. The CVCs from each production were digitally excised from their sentence context and presented randomly via headphones in a quiet lab. The listeners identified the vowel in each production by matching it to a labelled button on a custom *Praat* screen. Up to three replays were permitted. Although we attempted to elicit only 4 phonemic vowel targets from the speakers during the recording sessions, pre-screening of the productions indicated that several of them did not match any of the intended categories. To allow the judges as much flexibility as possible in identifying the productions, we therefore provided them with 10 vowel choices. In addition, an "other" category was available when no satisfactory match could be made. The judgment task was thus comparable in some respects to transcribing vowels by hand, except that the *Praat* presentation allowed automatic tallying of responses.

RESULTS

A variety of analyses were carried out to evaluate the effects of vowel category and consonantal context on vowel durations, and to pinpoint differences in patterning between the Japanese and Native English speakers. We focused on whether the Japanese speakers produced exaggerated duration differences between vowels, whether they used vowel duration to distinguish final voiced from voiceless consonants, and whether the intelligibility of their vowels was related to length. Where ANOVA and correlational analyses are reported, we used $p < .05$ as the criterion for significance. For all pairwise post hoc comparisons (*t*-tests), we used a Bonferroni adjustment to maintain an overall α of .05.

Vowel durations

We reduced the measurement data to obtain each speakers' mean duration for each rhyme by pooling over all the representations of that particular rhyme. Visual inspection of the data suggested a trend in both groups for tense vowels to be longer than lax vowels and for vowels to be longer before /d/ than before the voiceless consonants. However, statistical analyses were required to establish any between- and within-group differences in patterning.

For each L1 group, a two-factor repeated measures ANOVA was computed to probe the effects of vowel category (V, 4 levels) and coda consonant (CC, 3 levels) on duration. A significant two-way interaction between V and CC emerged for both the English speakers, [$F(6, 48) = 9.390$, $p <$

.0005], and the Japanese speakers, [$F(3.099, 71.268) = 4.470$, $p = .006$, Greenhouse Geisser adjustment due to violation of the sphericity assumption]. Follow-up single-factor ANOVAs yielded a significant effect of V for both speaker groups in all coda conditions; i.e., the Japanese and the English speakers both showed duration differences across vowels in all three consonant contexts (Table 2).

Table 2

Effects of Vowel identity on durations

Coda	L1 English	L1 Japanese
/t/	$F(3, 24) = 6.259$, $p = .003$	$F(3,69) = 69.777$, $p < .0005$
/k/	$F(3, 24) = 3.805$, $p = .023$	$F(3,69) = 58.344$, $p < .0005$
/d/	$F(3, 24) = 26.904$, $p < .0005$	$F(3,69) = 52.029$, $p < .0005$

Tense-lax pairs

Differences in tense-lax vowel durations were examined through pairwise within-group comparisons in the three coda conditions. As illustrated in Figures 1, 2, and 3, the Japanese speakers produced tense vowels with significantly longer durations than their lax counterparts regardless of coda. Effect sizes (Cohen's *d*) for the Vt and Vk pairs were large, ranging from 1.03 to 2.07 (Larson-Hall, 2015).

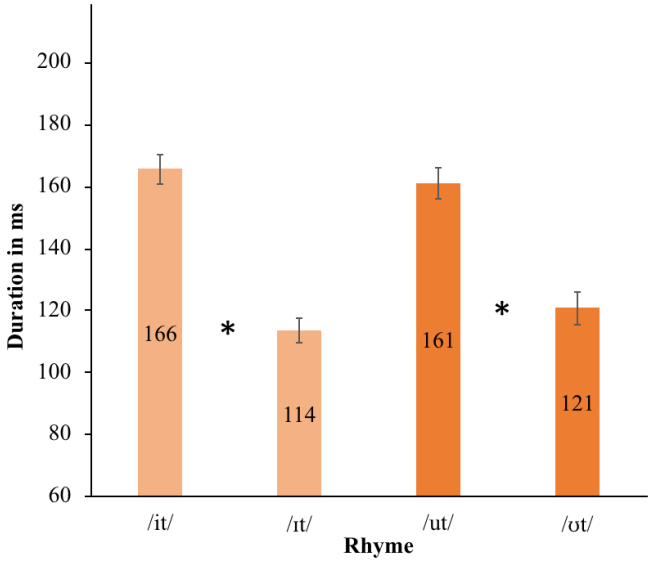


Figure 1: Japanese speakers' mean durations (with standard error) for the four vowels in /t/ context. (* $p < .05$)

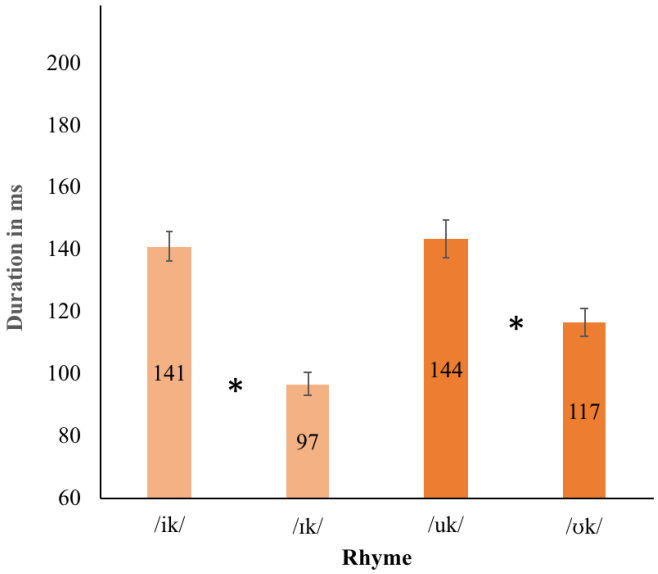


Figure 2: Japanese speakers' mean durations (with standard error) for the four vowels in /k/ context. (* $p < .05$)

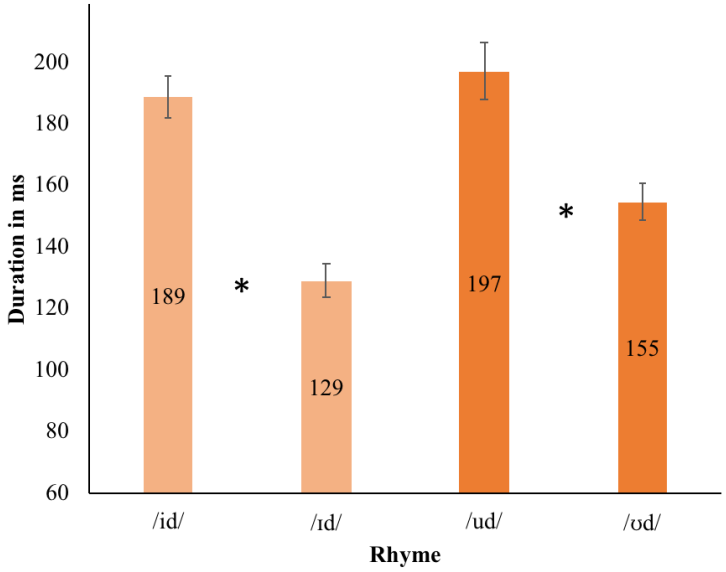


Figure 3: Japanese speakers' mean durations (with standard error) for the four vowels in /d/ context. (* $p < .05$)

The native English speakers, however, produced fewer significant differences. They produced /u/ significantly longer than /ʊ/ when followed by /t/ (Figure 4) and /i/ was longer than /ɪ/ when

followed by /d/. (Figure 5). None of the other differences, including those for vowels before /k/, reached significance. Effect sizes for Vt and Vk were noticeably smaller than for the Japanese group, with Cohen's *d* ranging from .41 to .91.

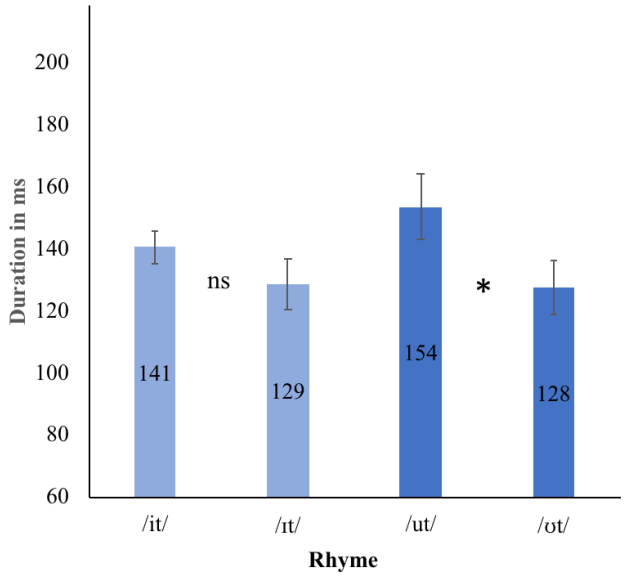


Figure 4: English speakers' mean durations (with standard error) for the four vowels in /t/ context. (* $p < .05$)

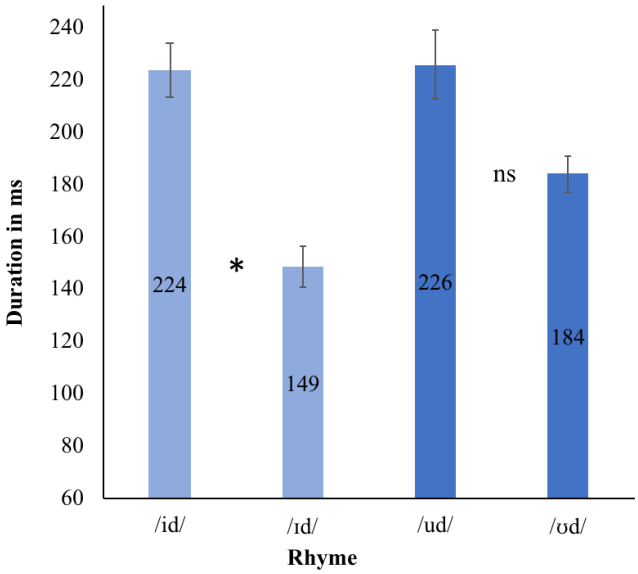


Figure 5: English speakers' mean durations (with standard error) for the four vowels in /d/ context. (* $p < .05$)

To further assess the degree to which the Japanese speakers exaggerated the tense/lax duration differences we computed mean lax/tense duration ratios for both speaker groups for the four sets of rhymes shown in Table 3. In all cases, the English speakers' ratios were comparatively closer to 1, indicating smaller duration differences between tense vowels and their lax counterparts than for the Japanese group. Pairwise independent samples *t*-tests confirmed that for all except the /ut-ʊt/ contrast, the Japanese speakers' ratios were significantly smaller than those of the native English speakers. To sum up, the Japanese group produced larger duration discrepancies in three of the pairs than did the English group, and there was a non-significant trend in the same direction for the fourth pair.

Table 3

Vowel duration ratios by rhyme for the two speaker groups

	Mean Lax/Tense Ratios	
Rhymes	Native English	Japanese
it, ɪt*	.92 (.161)	.69 (.103)
ik, ɪk*	.92 (.078)	.69 (.112)
ut, ʊt	.84 (.114)	.76 (.202)
uk, ʊk*	.96 (.067)	.83 (.133)

* $p < .05$

Effects of coda

Besides the effect of V, there were also significant effects of CC on the vowel durations of all four vowels for both the Japanese and English speakers, as shown in Table 3.

Table 3

Effects of coda on vowel durations

Vowel	L1 English	L1 Japanese
/i/	$F(1.23, 9.85)^{\dagger} = 69.452$, $p < .0005$	$F(1.45, 33.39)^{\dagger} = 51.590$, $p < .0005$
/ɪ/	$F(2, 16) = 6.970$, $p = .007$	$F(1.51, 34.65)^{\dagger} = 34.110$, $p < .0005$
/u/	$F(2, 16) = 33.923$, $p < .0005$	$F(1.29, 46) = 26.469$, $p < .0005$
/ʊ/	$F(2, 16) = 32.481$, $p < .0005$	$F(2, 46) = 44.440$, $p < .0005$

†Because of violations of the sphericity assumption, the Greenhouse-Geisser adjustment was applied.

For the sake of brevity, this portion of the paper focuses only on the effect of coda voicing on vowel durations; i.e., it compares only the V+/t/ durations with the V+/d/ durations. Pairwise within-group comparisons were run for both the Japanese and English speakers. As shown in Figures 6 and 7, the vowels before /d/ were generally longer than the matching vowels before /t/. In fact, almost all pairwise differences in both groups were significant, except for the lax vowel /ɪ/ as produced by the native English speakers.

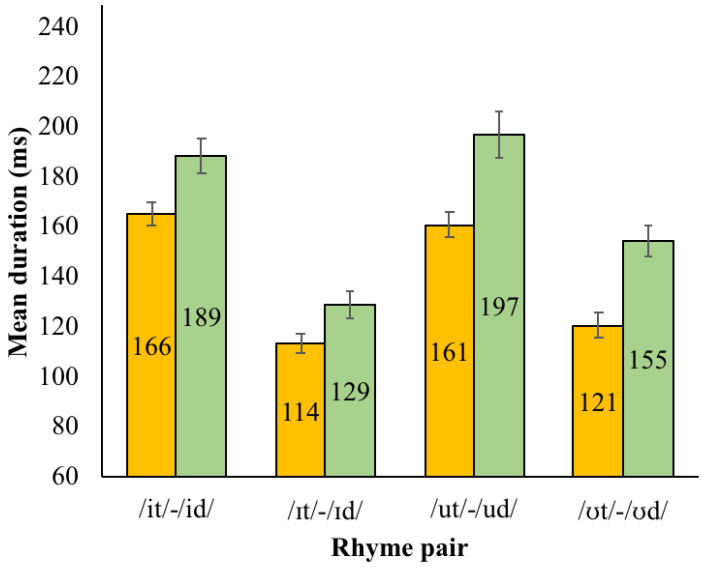


Figure 6: Effect of coda voicing on vowel durations for the Japanese speakers (standard error shown).

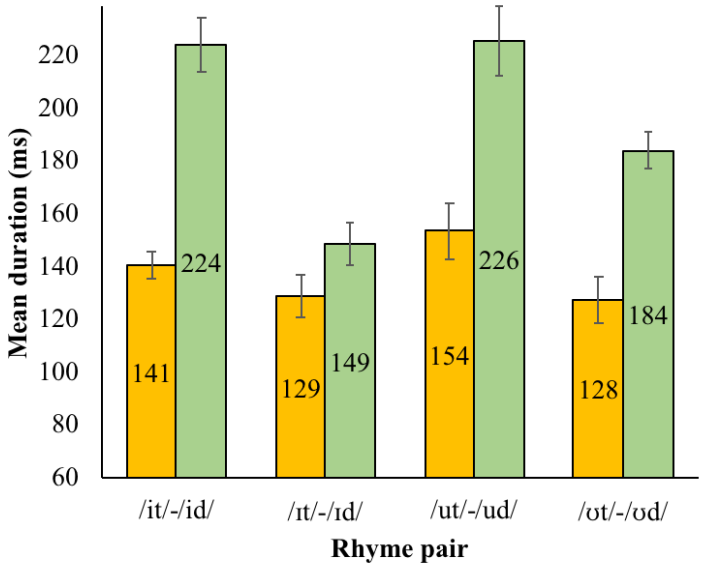


Figure 7: Effect of coda voicing on vowel durations for the Native English speakers (standard error shown).

Pairwise between-group comparisons on the duration ratios (voiceless/voiced) were computed to determine whether the Japanese speakers produced a native-like duration distinction in the vowels. The /it/-/id/ ratio proved significantly smaller for the native English group than for the Japanese group, indicating a greater discrepancy in duration on the part of the English speakers. The other between-group differences were not significant.

Vowel intelligibility

Vowel intelligibility was assessed on a correct/incorrect basis by tallying the number of times each judge identified a particular production as the actual target vowel. Using this criterion, an agreement rate of 83% was observed between the two judges, an outcome that compares very favourably with the results of other vowel intelligibility studies (e.g., Munro & Derwing, 2008).

When scores were pooled for the two judges, the overall intelligibility of the Japanese speakers' target vowels was 85%. However, several qualifications must be kept in mind regarding this number. First, there was noteworthy inter-talker variability in intelligibility: individual rates varied from 69% to almost 100%. Second, variability was found across vowels as well. The lax front vowel /ɪ/ was least intelligibly-produced (76%), while its tense counterpart /i/ was the most intelligible vowel of the four (93%). The back vowels /u/ and /ʊ/ showed intermediate intelligibility rates of 87% and 82%, respectively. Third, high overall intelligibility did not guarantee equally high performance on all four vowels. Speaker JVL09, for instance, performed considerably worse on /ɪ/ (69%) than on the vowels overall (87%). The reverse pattern was observed for JVL27, who performed relatively well on /ɪ/ (79%), even though her overall identification rate (69%) was the lowest of all speakers. Finally, vowel intelligibility also depended on the following consonant, with a lack of parallelism between the front and back vowels. While tense /i/ received higher identification rates than lax /ɪ/ when followed by /k/, the opposite tense-lax ordering was found for the back vowels.

Effect of vowel duration on intelligibility

The lax/tense vowel ratio data described earlier were used to assess the degree to which exaggerating the relative durational distinction between the tense and lax vowels would affect intelligibility. The Pearson *r* correlation between mean tense/lax ratios and overall vowel intelligibility was -.064 ($p > .1$), suggesting no meaningful relationship between the two. Furthermore, the three speakers who showed the most exaggerated tense/lax differences all scored above the median in terms of vowel intelligibility, one of them being the most intelligible of all. In summary, the data provided no evidence that exaggeration was predictive of reduced intelligibility.

DISCUSSION

It is well known that English tense vowels tend to be inherently longer than lax vowels (Crystal & House, 1988; House, 1961; Ueyama, 2000); however, the duration difference is relatively small and might not be readily noticed by English L2 learners. Nonetheless, the Japanese speakers in this investigation not only produced a significant duration difference in tense/lax pairs, but actually did so in more rhymes and to a greater degree than did the native English speakers. In other words, our expectation of exaggerated differences in duration on the part of the Japanese group (E1) was

upheld. The precise reasons for this exaggeration cannot be ascertained. On the one hand, the speakers may have assimilated the English tense vowels to their corresponding native Japanese long vowel categories and the lax vowels to their native short categories, perhaps because they had noticed duration differences in the English input that they received. On the other hand, it is also possible that they were taught to produce tense/lax differences as length differences, as is known to occur in some English language instruction (Wang & Munro, 2004).

A second expectation of ours (E2), namely that the Japanese speakers would not produce a native-like vowel duration difference to distinguish voiced from voiceless consonant codas received little support from our data. Lengthening before voiced obstruents has long been observed in English (House, 1961, p. 1177), but the dearth of possible consonant codas in Japanese means that no such effect occurs in the speakers' L1. Moreover, Tsukada's (2009) Japanese learners of English did not produce such a distinction. The fact that the Japanese speakers tended to display native-like lengthening of vowels before voiced stops in the present study suggests that they had learned to make use of this cue in spoken English. Although we cannot be certain of the reasons for this difference in findings, it is possible that Tsukada's participants had lower overall proficiency in English and had not yet had enough opportunity to acquire this length difference.

The one hypothesis tested in this study was that exaggeration of the durational distinction between tense and lax vowels would be linked to reduced vowel intelligibility. We proposed that speakers who relied mainly on duration to distinguish /i/-/ɪ/ and /u/-/ʊ/ would do so at the expense of producing accurate spectral differences. This prediction was not confirmed in the data. Not only was there no meaningful correlation between the extent of exaggeration and listeners' intelligibility assessments, but the speaker who showed the greatest degree of exaggeration was the one who produced the most intelligible vowels overall.

CONCLUSION

The findings reported here shed some additional light on the contributions of L1-to-L2 transfer and L2 acquisition processes to vowel production. First, the fact that the Japanese speakers exaggerated the durational differences between English tense and lax vowels suggests that they made use of their L1-based knowledge about vowel duration in their L2 implementations. Second, the data indicate that the Japanese group were successful at acquiring a new type of length distinction with no parallel in their L1.

Although this study was not designed with pronunciation teaching goals in mind, our findings are weakly suggestive of a few pedagogical consequences. In particular, the interspeaker variability that we observed in high vowel intelligibility (69% to nearly 100%) suggests that some learners may require more instruction than others in order to acquire the vowel distinctions at issue. Moreover, receiving instruction on the /i/-/ɪ/ distinction may be a worthwhile undertaking, since this contrast has a high functional load in English (see Tsukada, 2009). Yet our data showed that /ɪ/ was the least-intelligibly produced of all four vowels. We must be cautious in extrapolating from this study to classroom applications, however, since we did not look beyond vowel productions in a very limited speaking task; nor did we evaluate the benefits of any particular instructional techniques.

ACKNOWLEDGMENTS

We thank Amanda Huensch for many useful comments on an earlier draft of this manuscript, as well as Susan Morton and Trevor Wilson for assistance with data collection and speech measurements. We also greatly value the cooperation of our speakers and the judges.

ABOUT THE AUTHORS

Noortje de Weers (ndeweers@sfu.ca) is a PhD student under the supervision of Dr. Murray Munro at Simon Fraser University. She holds two master's degrees from the Netherlands, and worked as a secondary English teacher for two years before moving to Canada. Her main research interest lies in sociophonetics, specifically how a person's physical features can affect how their speech is perceived.

Murray J. Munro (mjmunro@sfu.ca) is a Professor of Linguistics at Simon Fraser University. His work on L2 pronunciation has appeared in a variety of international journals and edited volumes since the 1990s. His most recent books are Derwing, T. M and Munro, M. J. (2015). *Pronunciation Fundamentals: Evidence-based perspectives for L2 Teaching and Research.* (Benjamins) and Levis, J. M. and Munro, M. J. (2017). *Pronunciation*. (Routledge Critical Concepts Series).

REFERENCES

Bohn, O. S. (1995). Cross-language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 279-304. Baltimore: York Press.

Boersma, P., & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Retrieved May 2017 from http://www.praat.org/.

Crystal, T. H., & House, A.S. (1988). The duration of American-English vowels: An overview. *Journal of Phonetics, 16*(3), 263-284.

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

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America, 108*(6), 3013-3022.

House, A. S. (1961). On vowel duration in English. *Journal of the Acoustical Society of America, 33*, 1174-1178.

Ingram, J. C. L., & Park, S-G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, *25*, 343-370.

Larson-Hall, J. (2015). *A Guide to Doing Statistics in Second Language Research Using SPSS and R (2nd Edition)*. New York: Routledge.

Lee, B., Guion, S. G., & Harada, T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition, 28*(3), 487-513.

Liu, C., Jin, S-H., & Chen, C-T. (2014). Durations of American English vowels by native and non-native speakers: Acoustic analyses and perceptual effects. *Language and Speech, 57*(2), 238-253.

Munro, M.J. (1993). Production of English vowels by native speakers of Arabic: acoustic measurements and accentedness ratings. *Language and Speech, 36*, 39-66.

Munro, M. J. & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, *58*, 479-502.

Munro, M. J., 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.

Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., and Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics, 39*(2), 156-167.

Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America*, *51*, 1296-1303.

Tsukada, K. (1999). *An acoustic phonetic analysis of Japanese-accented English*. Unpublished doctoral dissertation, Macquarie University, Sydney, Australia.

Tsukada, K. (2009). Durational characteristics of English vowels produced by Japanese and Thai second language (L2) learners. *Australian Journal of Linguistics, 29*(2), 287-299.

Ueyama, M. (2000). *Prosodic transfer: An acoustic study of L2 English vs. L2 Japanese.* Unpublished doctoral dissertation, University of California.

Vance, T. J. (1987). *An introduction to Japanese phonology*. Albany: SUNY Press.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*, 539-552.

Garcia, C., Kolat, M. & Morgan, T. (2018). Self-correction of second language pronunciation via online, real-time, visual feedback. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 54-65). Ames, IA: Iowa State University.

SELF-CORRECTION OF SECOND-LANGUAGE PRONUNCIATION VIA ONLINE, REAL-TIME, VISUAL FEEDBACK

Christina Garcia, Saint Louis University

Mark Kolat, The Ohio State University

Terrell A. Morgan, The Ohio State University

We have built a set of web-based tools that take learning far beyond the classroom for L2 pronunciation students. Among these is a user-friendly site where students record themselves and can instantaneously compare their own voiceprint to that of a native speaker. Students continue to practice specific sounds until their production "looks" like their native target. For each sound, the system displays native and non-native speech samples (i.e., clickable sound files) alongside spectrographic representations. A column to the left orients the user on how to interpret the spectrograms to analyze the sound. The interface enables students to produce live-generated spectrograms, re-recording until their productions both sound and look like the native output. Recordings are automatically submitted to the instructor for effortless monitoring of students' progress. This interface was originally developed for Spanish and has been extended to English and Quechua. The response from Spanish students is overwhelmingly positive, many citing that "seeing" their pronunciation allowed them to internalize the steps necessary to make improvements as never before. They report increased satisfaction at being able to work on their pronunciation outside of class and receive the necessary feedback to assure that they are pronouncing correctly, even without a language expert present.

INTRODUCTION

In her keynote speech at this year's annual conference, Isabelle Darcy (2017) highlighted the ways in which, as researchers and teachers, we can bridge the gap between research and practice in pronunciation teaching. How do we apply our research to the pronunciation classroom and vice versa? One such gap between teaching and research can be seen in the way we have our students practice. Despite considerable technological advances in phonetics research in the past twenty years, we often find ourselves employing the same pronunciation drills we have always used rather than harnessing the technology used in modern acoustic analysis for pedagogical purposes.

Some scholars have noted this gap, experimented with different types of visual feedback, and found these methods to have enormous potential. In his review of Computer Assisted Pronunciation Training (CAPT) techniques, Levis (2007) supports the use of visual feedback in CAPT systems, citing the best-known visualizations as: spectrograms, waveforms, and pitch tracings. Studies specifically looking at the efficacy of CAPT programs have shown increased gains when compared to control groups for features such as pitch, voicing, vowel length, and

geminate consonants (Hew & Ohki 2004); vowel height (Quintana-Lara 2014, Kartushina et al. 2015); and rhotics (Patten & Edmonds 2015), among many other features.[1]

Thus far there has been a focus on learners of English, and as is evident in all areas of linguistic research, our pedagogy can be significantly improved by trying to solve the "problems" of other languages. Additionally, despite these advances, students often report technical frustration with programs such as *Praat*, and are unable to see beyond the technology itself in order to truly appreciate what it has to offer. Visual feedback can be powerful, but it seems that with today's students, how this feedback is packaged is of utmost importance.

The motivations for using technology in pronunciation practice are evident: face time with students is limited and instructors are constantly left wishing they had more time for individual feedback. While listening to recordings outside of class commonly supplements in-person feedback, it can quickly become overwhelming. On the student side, there are learners who are unable to hear the difference between their pronunciation and the native pronunciation, rendering it impossible to identify how to "fix" their own. Thus, pronunciation instructors are left with a few problems to solve: how can students receive effective and timely feedback without the instructor present? How can cutting-edge technology be efficiently used to simultaneously help students improve their pronunciation *and* learn about acoustic phonetics?

We have designed a user-friendly site[2] that aims to resolve these issues by allowing students to record themselves in real time and receive instantaneous feedback via spectrographic representation. Upon logging in to the interface, students are presented with native and non-native speech samples of a particular sound accompanied by spectrograms of these speech samples. They can listen to the difference between the native and non-native productions while simultaneously noting how the two spectrograms differ, guided by an instruction panel. Students are then able to record their own production and compare it to the native and non-native, both visually and auditorily. Students can continue to practice and re-record the sounds until their production 'looks' like that of the native target. Crucially, students do not need to download a program or native speech samples in order to use this tool; everything is provided in a one-stop shop so that the technical difficulties mentioned earlier are eliminated and students can focus on the visual feedback itself. In the sections that follow, we describe in detail the design of the tool and the user experience (section 2), feedback from students and instructors (section 3), the benefits and limitations of this methodology (section 4), and areas for its future development (section 5).

Description of the tool

The online tool that we have developed makes practicing difficult sounds for the L2 learner of Spanish, English and/or Quechua a smooth and efficient experience. For now, each sound is housed on a separate webpage and thus the students are provided with a list of web addresses for

[1] This list of studies is necessarily limited due to the length restrictions of this paper. For an overview of other studies in CAPT, see Levis (2007) and O'Brien (2006).
[2] Inspiration for this online interface grew out of our collaboration with the OhioSpeaks project of the Department of Linguistics at The Ohio State University (Wanjema et al. 2013).

all of the target sounds. Ohio State students utilizing this tool for a course are asked to sign in with their credentials so that their usage and progress can be monitored by their instructor.

Upon reaching the webpage of a unique sound or phonological process to be practiced, the user finds sidebar instructions on the left that are tailored to the particular lesson. The target sound and its non-native counterpart appear in phonetic notation, and the instructions are broken down into phases, the first being "Listen & Observe" and the second "Practice & Record". Next to these sidebar instructions, we find a series of three interactive voiceprint boxes where the spectrographic representations appear. The first and third boxes correspond to the Phase 1 instructions and contain prerecorded examples of the native (box 1) and non-native (box 3) pronunciations of the target sound. The student is able to listen to these recordings by clicking the "Play" button below each box and, at the same time, see a real-time spectrogram of the recordings in the boxes. The student can repeat Phase 1 as many times as they desire, being led by the instructions to understand what they are hearing and viewing. These instructions ask the student comprehension check questions that also contain descriptions of the spectrograms, presented in such a way as to highlight the differences between the native and non-native versions. For example, if the student is trying to perfect their pronunciation of the Spanish approximant [β], which would be realized as the full stop [b] in English, the page would appear as in Figure 1.

The spectrograms here clearly show that there is a large gap of white space between the colored columns of the non-native version where the red formant lines break, but there is more continuity in the red formant lines of the native version. While a written description of this comparison is contained in the Phase 1 instructions, students also receive basic instruction in interpreting spectrograms in class.

In Phase 2 the student is given the chance to practice their own pronunciation of the target sound by clicking the "Record" button under the box labeled "Your Version". Upon clicking the "Record" button, they do their best to imitate the native pronunciation, and a spectrographic representation of their speech shows up in real time as they speak. Contained in the Phase 2 part of the sidebar are explicit instructions on how to execute their own recordings as well as the English translation of the word that they are practicing that contains the target sound, in this case *haba* 'fava bean'.

Figure 1. Screenshot of online interface for Spanish [β]

Once the student has completed their first recording, the options to "Play" or go "Back to Carmen" (Ohio State's learning management system) appear, and the student also sees that their first attempt has been saved. The student can then play their recording in coordination with the native and non-native ones to auditorily compare them as well as visually compare the spectrograms. They are able to rerecord themselves as many times as they wish until they are satisfied with their native-like audio and visual realization of the target sound. At this point they click a "Save" button that appears under their spectrogram to save the recording. Both the automatically saved first recording and their elected final recording are sent to the instructor so that they are able to track the student's progress and identify strengths and weakness among all students in their class.

So far, the target Spanish sounds for which we have adapted this tool are: approximants [β] (vs. English [b]): *haba* (screenshots in Figure 1), [ð] (vs. English [d]): *hada,* and [ɣ] (vs. English [g]): *haga*; unaspirated voiceless stops [p] (vs. English [p^h]): *pan*, [t] (vs. English [t^h]): *tan*, and [k] (vs. English [k^h]): *can*; and fricative voicing [zð]: *desde* vs. *de este*.

For Quechua, thus far we have focused on the differences between the unaspirated, aspirated, and ejective voiceless stops and affricates: [p] vs. [p^h] vs. [p'']; [t] vs. [t^h] vs. [t'']: *tanta* ('meeting'), *thanta* ('old'), *t'anta* ('bread'); [k] vs. [k^h] vs. [k'']; [tʃ] vs. [$tʃ^h$] vs. [$tʃ^{\prime}$]. Figure 2 is a screenshot of the comparison of *t'anta* (box 1) vs. *tanta* (box 3). The target sound is the ejective [t''] that does not appear in Spanish or English, and here it is compared to the unaspirated version *tanta* which contains the [t] that is present (in this context) only in Spanish (a language already spoken by most learners of Quechua). If the student speaks English but not Spanish, comparing

t'anta (with the ejective) to *thanta* (with aspiration) would be more appropriate since the default English pronunciation would be aspirated.

In this example, the native version contains three distinctive segments with white space between each, the first of which being an initial, high-intensity explosion representing the ejective [t']. In the non-native version, this ejective does not appear because the word *tanta* is being produced, which simply contains an initial unaspirated stop followed immediately by the vowel [a]. There is therefore no white space between the stop [t] and the appearance of the red vowel formants.

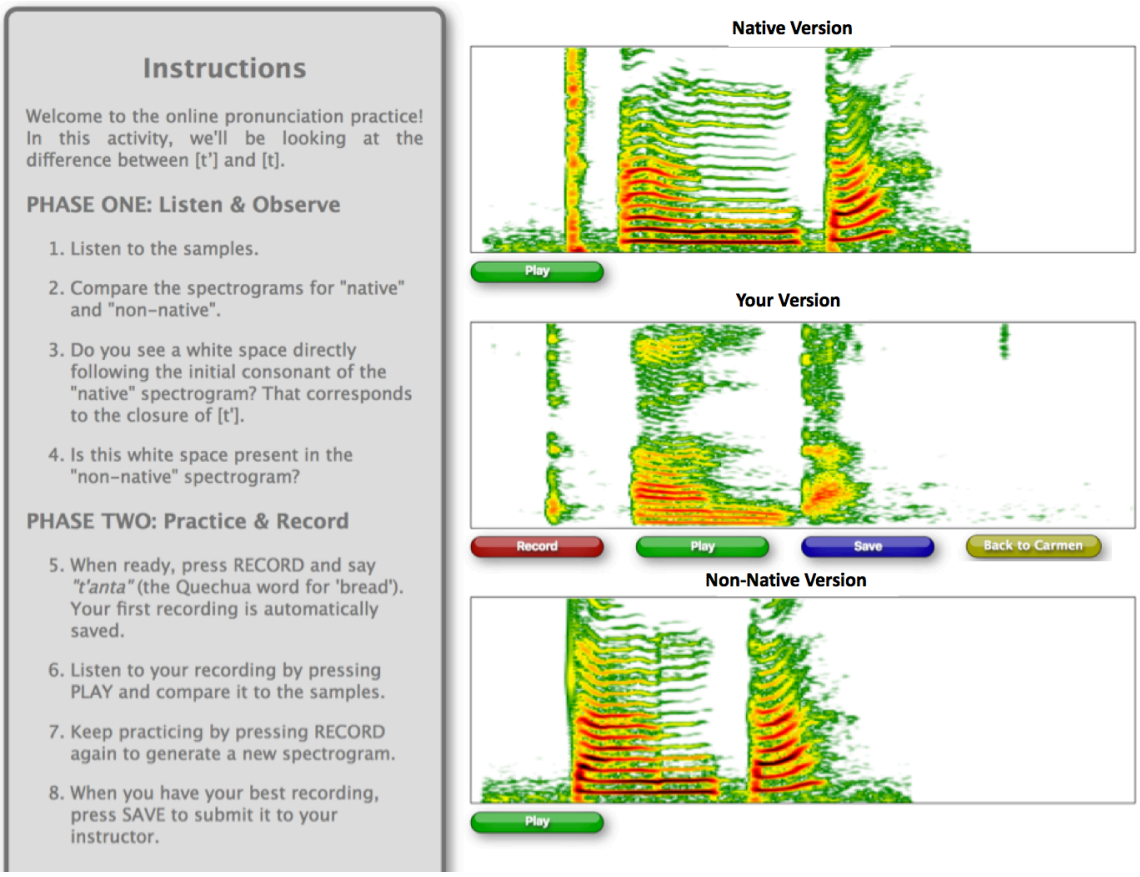


Figure 2. Screenshot of online interface comparing Quechua words *t'anta* 'bread' and *tanta* 'meeting'

Finally, for English, we have targeted the following: aspirated [p^h] (vs. [p] in many other languages: *pot*); aspirated [t^h] (vs. [t]): *top;* aspirated [k^h] (vs. [k]): *cop*; [h] vs. [ø]: *hall* vs. *all*; flapping /t, d/→[ɾ]: *kitty*, *buddy*; and bunched (postalveolar approximant) [ɻ̟]. Figure 3 displays a screenshot of [h] vs. [ø]: *hall* vs. *all*.

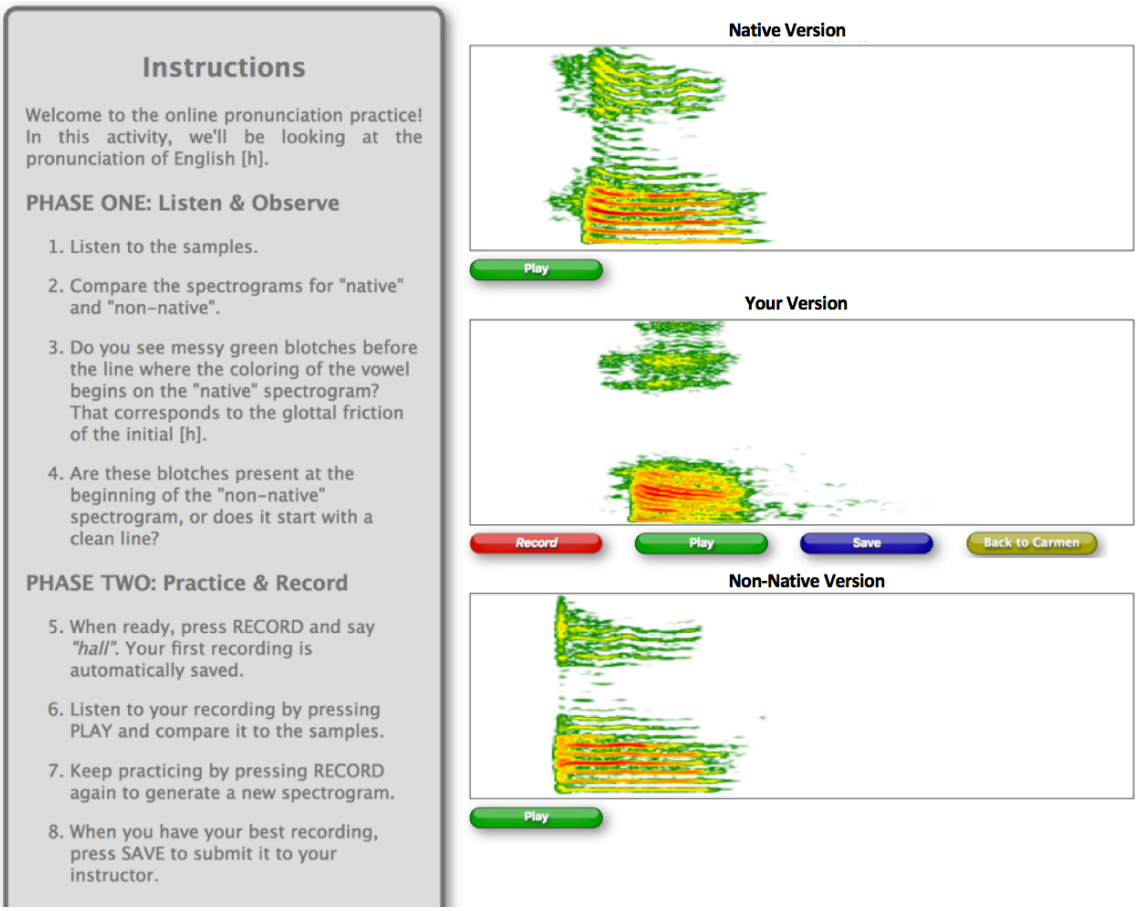


Figure 3. Screenshot of online interface comparing English words *hall* and *all*

In this example, the difference in frication is being focused on, with the native version containing frication of [h] as represented by the green friction before the yellow and red formants of the sound [a] begin. In the non-native version, there is no green friction before the yellow and red formants of the vowel [a] begin; it is just white. This represents the lack of the voiceless fricative [h], changing the lexical meaning of the word in English. This non-native version would be expected of students, for example, with L1s of French or Italian, which do not contain a voiceless glottal fricative in their phonetic inventory.

Having already implemented this tool into certain courses at Ohio State, we have received feedback from both students and instructors about their experiences. This feedback is outlined in the section that follows.

Student & Instructor Feedback

Thus far, this tool has been used in eight sections of Spanish 3404 (“Pronunciation”) at The Ohio State University since 2015. Quantitative survey results were collected in the first semester of its implementation (Autumn 2015), and since then qualitative impressions have been collected from both students and instructors. In this section we summarize this feedback, focusing on the areas most important for future development of the tool.

The quantitative survey results display the contrast between students' perceptions of online tools, in particular the pronunciation tool, at the beginning and end of the semester, before and after interacting with the actual tool. The survey included whether the online pronunciation activities *will be* (pre-survey, N=31) or *were* (post-survey, N=8) user-friendly, good preparation for class, worthwhile, enjoyable, and helpful in identifying and fixing pronunciation errors. At the beginning of the semester, students overall where somewhat skeptical about the utility of the online pronunciation tool, with rates of agreement vacillating between 45-55% agreement that it would be useful, worthwhile, helpful, etc. In the post survey, these agreement rates raised to 63-75% in the qualities of: user-friendly, easy to follow, enjoyable, teaching needed skills, helping to prepare for class, and identifying pronunciation errors.

In qualitative evaluations, students are quick to point out the advantage of receiving immediate feedback on their pronunciation. Pronunciation students and instructors alike are often frustrated with the small amount of class time in which feedback can be given. The tool by no means replaces invaluable instructor feedback; however, it offers students an *additional* way to receive quality feedback, without having to wait until the next class period or for their instructor to grade their recording. Instructors have recognized that this feature increases students' autonomy and sense of responsibility for their pronunciation gains. Interaction with the tool affords students a new sense of linguistic awareness that aids in meeting their pronunciation goals.

Students also repeatedly report that they enjoyed and appreciated the visual aspect of the tool's feedback. They like being able to see their voice visualized in a 'techy' way they have never seen before, especially those who are visual learners. Speaking to the need for multimodal learning (cf. Levis 2007: 14), several of these visual learners noted that, in class, they had difficulties hearing the difference between their own pronunciation and the native model of their instructor, but that the tool allowed them to understand and internalize their own pronunciation in a new way by giving them a visual target to work towards. More specifically, the fact that the interface includes both native Spanish and naïve native English productions gives students an idea of the two extremes on the spectrum and, even if they do not succeed in 'matching' the native, they might land somewhere in between. From the instructor's perspective, the visualization reinforces the Spanish versus English contrasts discussed in class and helps students to better conceptualize the phonological system. To put it succinctly, "seeing is believing" and this tool can convince students that the contrasts we talk about in class actually exist. One instructor also noted a feedback loop in which seeing the linguistic contrasts actually caused his students to be able to hear these contrasts better moving forward.

Finally, students like the interactive and creative aspect of the pronunciation practice. A couple of instructors had their students 'play around' with a super 'gringo' accent versus their best native accent, which encourages students to do the activity over and over again and see how the visual changes. When used properly, this tool can take one of the most tedious parts of language acquisition (pronunciation drills) and make it more fun. This feature of the system leaves students wanting more, which encourages them to continue taking linguistics, and specifically phonetics, as they further their studies.

Comparisons with other technologies

The online tool described here is just one of a number of emerging technologies with implications for L2 pronunciation instruction. Like all computer-assisted methodologies, it presents both virtues and limitations, which we will examine here in more detail.

To begin with, some articulatory distinctions can be clearly seen on a spectrogram that cannot be appreciated using other visual representations, and vice versa. While electropalatography illustrates nuances in the contact between articulators which are not easily discernable on a spectrogram, it is only able to do so if, in the production of the sound in question, the tongue makes contact with some part of the hard palate. Vowels and labial consonants (just to name two classes of sounds) are therefore irrelevant to the usefulness of electropalatograms as a teaching tool.

Similarly, the distinction between velar and uvular consonants (critical in Bolivian Quechua) can be illustrated neatly via ultrasound technology (Bird & Bliss 2017), but not so clearly on a spectrographic voiceprint. The spectrogram, however, is able to capture labials and other sounds whose articulators interact outside of the section of the vocal tract captured in the ultrasound image. And while relative tongue height and backness are visible in the formants on a spectrogram, the nuances of vowel quality are much clearer on a vowel plot, such as that produced by the computer program described in Lie-Lahuerta (2014).

Some studies have critiqued the use of spectrograms in pronunciation practice (cf. Neri, Cucchiarini, & Strik 2002), alleging that students are not likely to be able to pick up on the errors they are making via such feedback. However, we have shown that with a small amount of training in interpretation and an interface that allows for native vs. non-native comparison, the use of spectrograms can be beneficial to L2 pronunciation students. What is key is making sure the spectrographic representations are used in a motivated and targeted way and highlight particular features that are relatively easy to observe in this type of visualization.

In fact, one additional benefit that accrues to the learner is basic knowledge of acoustic phonetics—for example, how approximants appear different from stops, or the relationship between, say, aspiration or glottalization and the subsequent onset of voicing. These concepts, and their articulatory concomitants, are neither trivial nor impractical in the acquisition of native-like pronunciation, and it therefore makes sense to teach them in conjunction with the targeted practice in question. There is even reason to believe that students will better internalize such notions when presented on a need-to-know basis and as part of an interactive lesson with personalized impact.

Areas of future development

While student and instructor feedback has shown the usefulness of this pronunciation tool, there are of course ways in which is can be improved, especially at a time in which the advances in CAPT technology multiply by the day. The first area of development that we see regards content as opposed to technical features. We continue to work on the expansion of this tool to include Quechua and English pages, with the eventual goal of expanding to other world languages. This understandably requires the collaboration of instructors of other languages, but presents vast potential for several reasons, one being that the addition of new languages and sounds challenges the limits of the online interface, ideally making it more powerful and flexible. With the development of pages in other languages, the system should be flexible enough to allow for

different comparisons based on the student's L1, as mentioned above in the case of Quechua *t'anta*. Another area of content revision concerns which sounds are targeted. Currently the pronunciation tool targets those linguistic features (ex. aspiration) that are easiest to observe on a spectrogram; however, additional attention should be given to whether these are the sounds that are most difficult for L2 learners and whether there is a way to expand the system to other 'difficult' sounds. Finally, we intend to expand the database of native speaker recordings to allow for the students to choose the dialect and gender of the native speaker that their speech is compared to, given that recent developments in the 'golden speaker' technology have shown that learners benefit greatly from hearing a model voice that is similar to their own (Ding et al. 2017).

Beyond content development, there are several areas in which the tool itself can be improved in order to allow for an even more positive student experience. The quality of the spectrograms, although somewhat dependent on the students' recording environment, should be improved to facilitate the identification of visual differences between native and non-native spectrograms. Also, we have received several comments about having some instantaneous feedback beyond the visual comparison of spectrograms. For instance, for the aspiration of /p, t, k/, the interface could let users know what their VOT (Voice Onset Time) was in order to compare to the VOTs of the native/non-native samples. This would require the use of a *Praat* (Boersma & Weenink 2017) script or similar and potentially more linguistic training for the users; however, for some sounds it would resolve the issue of users not knowing what to look for in the spectrogram.

To that point, the main area of future development that we see at this time is figuring out how to make the interface even more user-friendly so that it could be used with novice and intermediate learners, who may or may not be receiving explicit pronunciation instruction in class. We feel that spectrographic representations have the potential to be used in this way if some adaptations were to be made. For instance, in addition to the instructions on the left side panel, there could be boxes or arrows that highlight the portion of the spectrogram that is relevant to the target sounds. There could also be a training module (cf. Levis 2007) added in order to guide the user through the ideal experience. Nevertheless, the question remains of how a novice or intermediate learner would apply the spectrographic information to their pronunciation. This presents a significant challenge, but one that might be overcome through the use of pronunciation podcasts to accompany the pronunciation interface.

Finally, we are currently working on making this site open to the public so that instructors and students at other institutions may take advantage of it. Ideally this version of the site will have a landing page that will allow users to peruse all of the sounds available for practice. It is our hope that by making the site open we are able to receive increased feedback that will allow us to continue to improve the tool's effectiveness. Look for it in the future at *pronounce.osu.edu*.

ACKNOWLEDGMENTS

We are grateful for the financial support of The Ohio State University's Impact Grant, and most especially the on-going and invaluable work of Henry Griffy and the technical expertise of Justin Slauson and Abhijit Varde. The student feedback reported here was collected under IRB protocol 2015B0261.

ABOUT THE AUTHORS

Christina García, Assistant Professor of Spanish and Linguistics at Saint Louis University, is a sociolinguist and phonetician interested in why Spanish speakers talk the way they do, and how to get students excited about the linguistic diversity of the Spanish-speaking world. She has studied and done fieldwork in Argentina and Ecuador, and her research examines how sounds are socially meaningful and contribute to the formation of regional identities. Christina bring her research into the classroom by challenging pronunciation students to use the same techniques she uses in her research (acoustic analysis) to internalize and improve their own pronunciation in new and interactive ways.

Christina García, Ph.D.
Assistant Professor of Spanish & Linguistics
Saint Louis University
1513 Morrissey Hall
3700 Lindell Blvd.
St. Louis, MO 63108
Phone: 513.505.2494
Email: christina.garcia@slu.edu

Mark Kolat is a first-year MA/PhD student in Hispanic Linguistics at The Ohio State University. He holds Bachelor degrees in Spanish and French from his present university, and has studied in Spain and Senegal. With previous studies of Spanish, French, Italian, Wolof, Swahili, and Quechua, his research interests include sociolinguistics, phonetics & phonology, and second language acquisition, particularly related to less commonly spoken/studied languages of Africa and Latin America. In serving as an L2 English/Spanish instructor in Senegal as well as spending a semester working to promote the Quechua language program at Ohio State, Mark hopes to expand the current applications of this pronunciation tool to other languages that he encounters through his graduate work.

Mark Kolat
MA/PhD student
Hispanic Linguistics
The Ohio State University
298 Hagerty Hall
1775 College Road
Columbus, OH 43210-1340 USA
Phone: 440.602.2781
Email: kolat.5@osu.edu

Terrell A. Morgan, Professor of Hispanic Linguistics at The Ohio State University, is a phonologist and dialectologist concerned with documenting linguistic diversity and finding new ways to put students, teachers, and fellow researchers in touch with the intimate details of the sounds and structures of Spanish. He has taught and done field work in Spain, Mexico, Guatemala, Honduras, Nicaragua, Costa Rica, Panama, the Dominican Republic, Puerto Rico, Colombia, Ecuador, Peru, Chile, Bolivia, Paraguay, Argentina, and Uruguay, as well as in

Hispanic communities in the United States. His Spanish phonetics textbook, *Sonidos en contexto: Una introducción a la fonética del español con especial referencia a la vida real*, was published in 2010 by Yale University Press.

Terrell A. Morgan, Ph.D.
Spanish & Portuguese
The Ohio State University
298 Hagerty Hall
1775 College Road
Columbus, OH 43210-1340 USA
Phone: 614.292.9555
Email: morgan.3@osu.edu

REFERENCES

Bird, S. & H. Bliss. (2017). Visualizing speech in a classroom setting using interactive ultrasound imaging. Paper presented at *Pronunciation in Second Language Learning & Teaching (PSLLT) 9th Annual Conference*, University of Utah.

Boersma, P. & Weenink, D. (2017). Praat: doing phonetics by computer [Computer program]. Version 6.0.35, retrieved 16 October 2017 from http://www.praat.org/

Darcy, I. (2017). "'A psycholinguist walks into a classroom…': A road-map for bridging research and practice in pronunciation teaching". Plenary talk at *Pronunciation in Second Language Learning & Teaching (PSLLT) 9th Annual Conference*, University of Utah.

Ding, S., C. Liberatore, G. Zhao, S. Sonsaat, E. Chukharev-Hudilainen, & Levis, J. (2017). Golden Speaker Builder: An interactive online tool for L2 learners to build pronunciation models. Paper presented at *Pronunciation in Second Language Learning & Teaching (PSLLT) 9th Annual Conference*, University of Utah.

Hew, S. H., & Ohki, M. (2004). Effect of animated graphic annotations and immediate visual feedback in aiding Japanese pronunciation learning: A comparative study. *CALICO Journal, 21*(2), 397-419.

Katrushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America, 138*(2), 817-832.

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

Lie-Lahuerta, C. (2014). Fix your vowels: Computer assisted pronunciation training. Retrieved from Holt, D. E. & Lord, G., *IGNITE: CASPSLaP resources.* www.ignite-caspslap-2014.weebly.com, Retrieved 29 October 2017.

Morgan, T. A., & García, C. (2016). Seeing is believing: Student self-correction of Spanish pronunciation via online, real-time, visual feedback. Retrieved from Campos-Astorkiza, Rebeka, Terrell A. Morgan, and Holly Nibert. *IGNITE: CASPSLaP Resources.* Online: u.osu.edu/ignite2016/.

Neri, A., C. Cucchiarini, & H. Strik. (2002). Feedback in computer assisted pronunciation training: When technology meets pedagogy. *Proceedings of "CALL professionals and the future of CALL research" Conference* (pp. 179–188). Antwerp, Belgium. Retrieved October 27, 2017 from http://repository.ubn.ru.nl/bitstream/handle/2066/76206/76206.pdf

O'Brien, M. (2006). Teaching pronunciation and intonation with computer technology. In L. Ducate & N. Arnold (Eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching* (CALICO Monograph Series, Vol. 5, pp. 127–148). San Marcos, TX: CALICO.

Patten, I., & Edmonds, L. A. (2015). Effect of training Japanese L1 speakers in the production of American English/r/using spectrographic visual feedback. *Computer Assisted Language Learning, 28*(3), 241-259.

Quintana-Lara, M. (2014). Effect of Acoustic Spectrographic Instruction on production of English /i/ and /I/ by Spanish pre-service English teachers. Computer Assisted *Language Learning, 27*(3), 207-227.

Wanjema, Shontael, Katie Carmichael, Abby Walker, & Kathryn Campbell-Kibler. (2013). Integrating teaching and large-scale research through instructional modules. *American Speech*, *88*(2), 223-235.

Liu, D. (2018). Translanguaging in prosody teaching: Beyond monolingual ideology. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 66-75). Ames, IA: Iowa State University.

TRANSLANGUAGING IN PROSODY TEACHING: BEYOND MONOLINGUAL IDEOLOGY

Di Liu, Boston University

Prosody plays an essential role in pronunciation teaching (Anderson-Hsieh, Johnson, & Koehler, 1992; Derwing, Munro, & Wiebe, 1998). However, some L2 English speakers do not use English prosody effectively (Pickering, 2001, Wennerstrom, 1998). In recent years, a number of studies have argued for similarities between the pragmatic functions of Mandarin and English prosody (Chen & Gussenhoven, 2008; Ouyang & Kaiser, 2015), suggesting the possibility of leveraging learners' L1 into prosody teaching. However, research studies investigating the efficacy of cross-linguistic prosody pedagogy are lacking. This study investigates the efficacy of a monolingual (English) metalinguistic awareness enhancement based prosody teaching method (mono-MAET) and a translingual (English and Mandarin) metalinguistic awareness enhancement based prosody teaching method (trans-MAET) by analyzing the pitch height of the sentence stress in L2 English speakers' read aloud speech. The participants who received trans-MAET demonstrated a statistically significant increase in the pitch height of the stressed constituents. This study informs teaching by showing that tapping into similar pragmatic functions across languages can lead to improvement.

INTRODUCTION

Prosody, also commonly referred to as suprasegmentals, includes a variety of speech features including intonation, rhythm, and stress. In the past 30 years, researchers have found that prosody plays a crucial role in pronunciation teaching and learning (Anderson-Hsieh, Johnson, & Koehler, 1992; Celce-Murcia, Brinton, & Goodwin, 2010; Derwing, Munro, & Wiebe, 1998). However, some L2 English speakers fail to exploit English prosody (Pickering, 2001; Wennerstrom, 1998) in communicating meaning, which may lead to some issues. For instance, L2 English speakers may not be able to actively participate in classroom discussions, and students who are native speakers of English may find it difficult to follow and understand the speech of their international teaching assistants.

However, the techniques used in prosody teaching are far from optimal. Reed and Michaud (2015) concluded that classroom pronunciation teaching, including prosody teaching, focuses mostly on imitation, drills, and repetition, which, according to Gilbert (2014), is not ideal as learners do not typically apply what they learned in the classroom in everyday communication. To solve this issue, Reed and Michaud (2015) argued that metalinguistic awareness of the importance and functions of prosody is essential in prosody teaching. However, studies that have investigated the efficacy of metalinguistic awareness based prosody teaching methods are still needed.

Raising learners' metalinguistic awareness of the importance and functions of English prosody, however, may not be sufficient to foster learners' use of English prosody due to the many ways that prosody can affect meaning in discourse. For instance, knowing that a change of sentence

stress can lead to a change in sentence meaning but being uncertain of the word(s) that should be stressed in different sentences, learners may be reluctant to use prosodic patterns because they do not want to be misinterpreted by native speakers of English.

Some researchers suggest that, "productive future avenues of research might involve investigations of the role of cross-language similarity in the learning of prosodic features" (Trofimovich, Kennedy, & Foote, 2015, p. 357). For example, Cruz-Ferreira (1987) talked about positive transfer and stated that, "Listeners are able to refer to general intuitions about the more likely meaning associated with lower and higher pitch." (p.116). This cross-linguistic approach has some advantages. On the one hand, L2 English speakers may no longer consider English prosody "decorative" if they know that they have been using prosody in a similar manner in their L1s. On the other hand, they may feel more confident in using English prosody if they realize that they can positively transfer some prosodic functions/usages from their L1s to English. Research studies investigating and comparing Mandarin and English prosody reveal that there are some similarities between the prosodic features and functions of these two languages (Chen & Gussenhoven, 2008; Ip & Cutler, 2016), suggesting the possibility of leveraging the L1 in Mandarin speakers' English prosody learning. However, research studies surveying the efficacy of cross-linguistic prosody teaching method are still lacking.

Translanguaging approaches assume that multilingual speakers strategically employ different semiotic systems to realize similar functions, supporting the potential advantages of cross-linguistic based prosody teaching. First developed in Welsh by Williams (1996), translanguaging was used as an approach to bilingual education. It was then expanded by the researchers as a theory explaining how bilinguals manage the linguistic resources in their cognitive system (Gracia & Li, 2014; Li, 2011). Mazak (2017) stated that translanguaging "posits that bilinguals do not separate their 'languages' into discrete systems, but rather possess one integrated repertoire of languaging practices from which they draw as they navigate their everyday bilingual worlds" (p. 5). Mazak (2017) further argued that translanguaging is "a pedagogical stance that teachers and students take on that allows them to draw on all of their linguistic and semiotic resources as they teach and learn both language and content material in classrooms" (p. 5).

To address current issues of prosody teaching and to investigate the role of metalinguistic awareness and translanguaging in prosody teaching, this study surveys the efficacy of two prosody teaching methods for the English sentence stress learning of Chinese L2 English speakers. Sentence stress is studied for two reasons. First, variations in sentence stress can shift the meaning or implication of a sentence. For example, in the sentence "She made the wrong decision," if the speaker stresses the word "she", the speaker emphasizes the fact that a particular person, not someone else, made the wrong decision. However, if the speaker stresses the word "wrong," the speaker emphasizes that the decision is not a right decision. Sentence stress is studied also because of a discrepancy in Chinese L2 English speakers' English and Mandarin sentence stress production. On the one hand, researchers have found that Chinese L2 English speakers' English speech is characterized by a flat pitch contour, suggesting the lack of sentence stress (Pickering, 2001). On the other hand, researchers have found that, like English, Mandarin uses sentence stress to signal information structure (Ip & Cutler, 2016).

This study compares a monolingual (English) metalinguistic-awareness enhancement based prosody teaching method (mono-MAET) and an extended translingual (English and Mandarin)

metalinguistic-awareness enhancement based prosody teaching method (trans-MAET) to a control group. The research questions are:

1. How does mono-MAET influence Mandarin speaking L2 English speakers' English prosody production, as measured by the pitch height of sentence stress?
2. How does trans-MAET influence Mandarin speaking L2 English speakers' English prosody production, as measured by the pitch height of sentence stress?

METHOD

This study adopts a pretest, intervention, and posttest design. Fifteen Mandarin-speaking L2 English speakers were randomly assigned to three groups (N=5). In the pretest, the participants were asked to read aloud a passage in front of a computer at a soundproof booth. The passage (see Appendix) was adapted from Hahn (2004). Participants were recorded using a voice recorder Zoom H4N. Before the participants read the passage aloud, they were asked to read through the passage silently, familiarize themselves with the content, and ask the researcher any questions they had. In the intervention phase, the participants were given different treatments. Participants in group 1 received the mono-MAET in English, participants in group 2 received the trans-MAET in both English and Mandarin, and participants in group 3 (the control group) answered interview questions. The intervention phase lasted for approximately 20 minutes. After the intervention, all participants were asked to read aloud the same lecture again.

The instruction materials for group 1 (mono-MAET) and group 2 (trans-MAET) were adapted from Reed & Michaud (2005). The materials were pre-recorded as two lectures by a native speaker of English and a native speaker of Mandarin. The trans-MAET lecture included the content of the mono-MAET and had additional examples in Mandarin, which were the literal translation of the English examples. The lectures were played once to the participants using computer software Praat (Boersma, 2001). The participants were allowed to ask questions while listening to the lectures.

The mono-MAET contained four parts: introduction, diagnostic, analysis, and practice. The concept of English sentence stress was introduced first, followed by a diagnostic task testing participants' ability to identify stressed constituents in a dialogue. The stressed constituents in the diagnostic task dialogue were then analyzed. Finally, the participants were asked to practice the use of English sentence stress using sample sentences provided.

In the introduction phase, the participants were given a brief introduction of English sentence stress. They were told that stressed words in a sentence are longer, louder, and higher in pitch. The participants were then asked to listen to the following set of sentences and identify the stressed word in each sentence.

1. The *teacher* didn't grade your exam.
2. The teacher didn't grade *your* exam.
3. The teacher didn't grade your *exam*.

In the diagnostic phase, the participants were asked to listen to the conversation below and underline the words that they thought should be stressed. The participants then listened to two

native English speakers having the same conversation and compared the words they underlined to the words that the two native English speakers stressed.

A: "Should I get the shirt with buttons or without buttons?"

B: "How about with the buttons?"

A: "Well, should I get the short-sleeved one or the long-sleeved?"

B: "Well, I like short-sleeved shirts."

A: "Okay, do you like the blue or the white?"

B: "I like the yellow. Can we go now?"

In the analysis part, participants listened to a paragraph explaining the concept of sentence stress using the examples from the conversation. For example, the participants were told that, when a speaker says, "I like short-sleeved shirts." with standard stress, there's no extra meaning and we don't know anything about the context. The sentence is neutral. However, when a speaker says, "I like *short*-sleeved shirts" with extra stress on the word "short," the speaker is making a contrast between *short*-sleeved and *long*-sleeved shirts. There is extra meaning and the sentence is contrastive.

Finally, the participants were asked to practice the use of sentence stress using some sentences provided (e.g., Yesterday we discussed the creation of Facebook. Today we'll discuss the marketing of Facebook).

The trans-MAET not only contained the content of the mono-MAET but also compared Mandarin and English sentence stress by providing literal translation of all the examples in the mono-MAET.

RESULTS

Participants' speech was analyzed using the speech analysis software Praat. The analysis focused on the pitch height of 18 contrastive stress words in the passage: "individualism, personal-1, group-1, collectivism, group-2, personal-2, coworkers, longer, you, shorter, collectivist, give in, individualist, go against, upper, lower, top, bottom".

Cruz-Ferreira (1987) found that L2 speakers use "the pitch height strategy" as an interpretive strategy for intonational meaning. She stated that "where the meaning contrast conveyed by intonation in L2 can be associated with broadly similar uses of pitch contours or pitch height in L1, 'abstract' generalizations regarding meaning seem to be made…" (p. 115). In this study, pitch height was also used as an indicator of the participants' use of sentence stress. Because the minimum pitch level measured by computer software might be skewed due to the existence of creaky voice, pitch height in this study was measured by subtracting the average pitch level of the whole passage from the maximum pitch level of the stressed constituents (see Figure 1). For each word analyzed, the maximum and average pitch level (in semitones) was obtained using the built-in pitch elicitation function of Praat. To make comparisons between speakers of different genders and ages, the pitch of the stressed words was analyzed in semitones relative to 100 Hz.

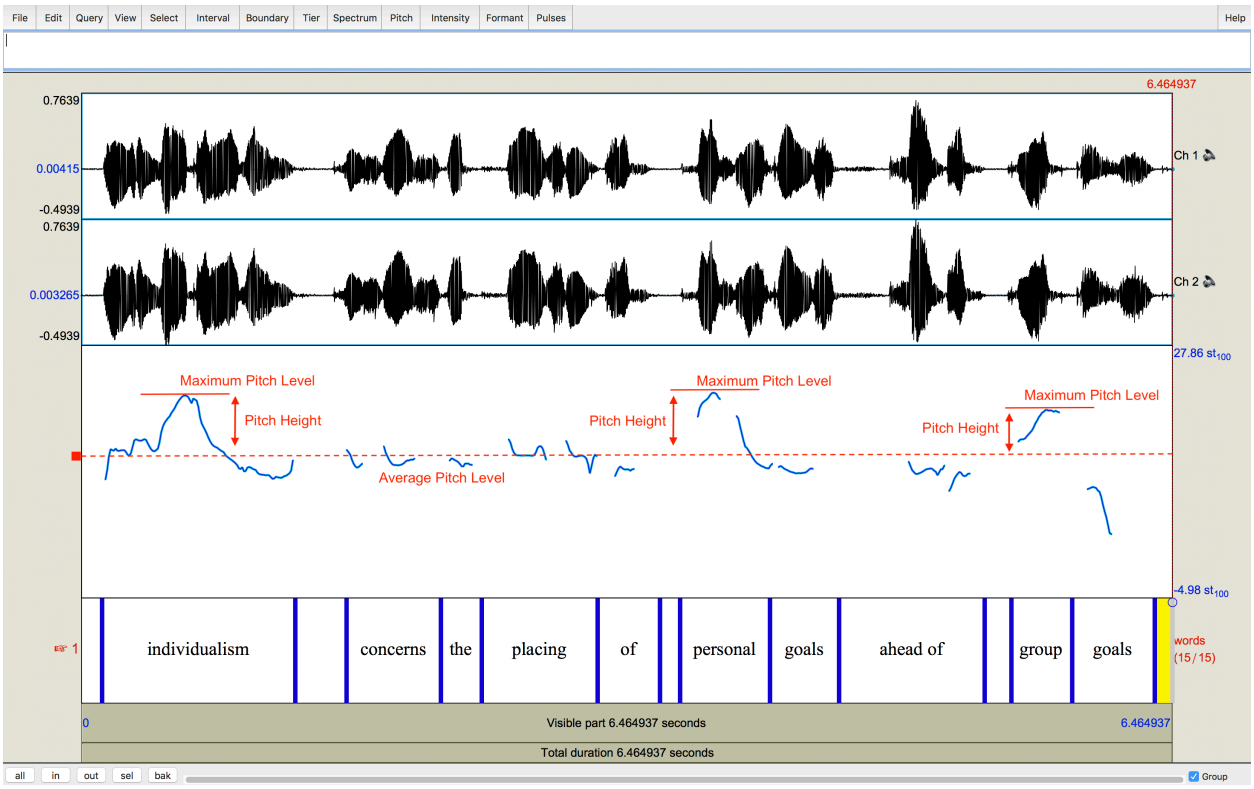


Figure 1. Pitch Height Measurement

The average pitch height of the stressed constituents in the participants' pretest and posttest speech was compared using the data analysis function of the software R (R Core Team, 2017). The average pitch height of the stressed constituents in Group 1 and Group 2 participants' speech both increased after the intervention. The pitch height of Group 2 (trans-MAET) had the biggest increase. The pitch height of the stressed constituents in Group 3 (control group), however, slightly decreased after the intervention (see Figure 2). Based on Pickering (2001)'s finding that Mandarin speaking L2 English speakers tend to use a "flat" pitch contour, the increase of the pitch height of the stressed constituents is considered as an enhancement of participants' ability to signal sentence stress.

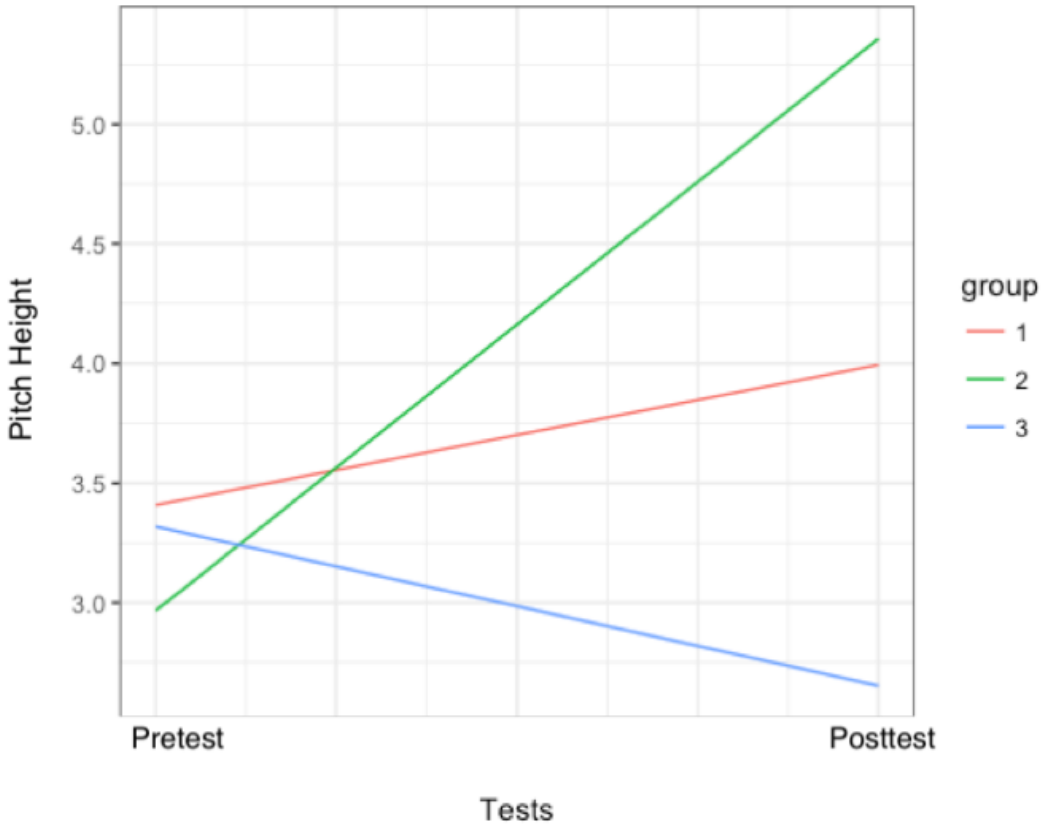


Figure 2. Average Pitch height of the stressed constituents

Pitch height of the stressed constituents was then analyzed using the linear mixed-effects model. The results show that the increase of the pitch height of the participants in Group 2 (trans-MAET) was statistically significant (t=7.474, df=4, p=.0017). The increase of the pitch height of the participants in Group 1 (mono-MAET) was not statistically significant (t=2.262, df=4, p=.0865). The decrease of the pitch height of the participants in Group 3 (control group) was not statistically significant (t=-2.205, df=4, p=.0921). (see Figure 3).

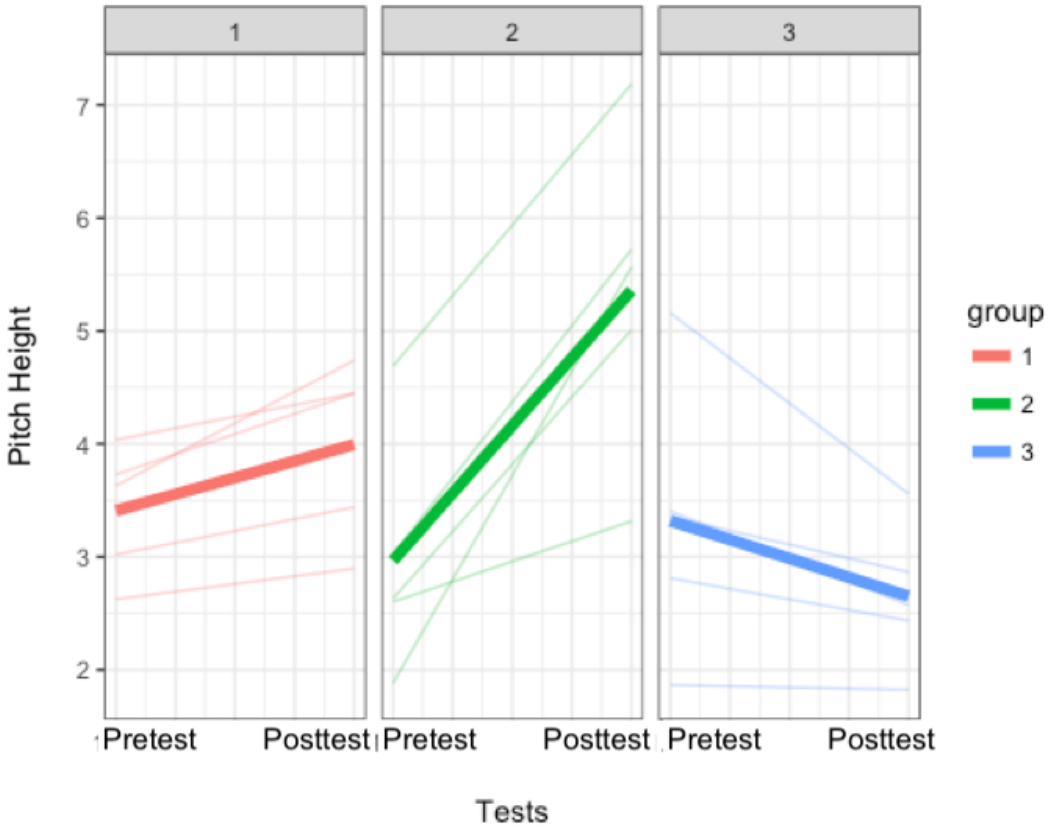


Figure 3. Linear regression of the pitch height of the stressed words.

DISCUSSION AND CONCLUSION

Some researchers suggested that Mandarin's tone system may interfere with the learning of English prosody for L2 English speakers (Clennell, 1997). However, this study finds that prosody instruction comparing Mandarin and English sentence stress (i.e. trans-MAET) increased the pitch height of Mandarin-speaking L2 English speakers' stressed constituents. This result suggests that Mandarin has some prosodic functions that can be transferred to English, which may facilitate learners' English prosody learning.

Reed and Michaud (2015) argued that metalinguistic awareness is the key to prosody teaching. This study supports their argument by showing that even a 20 minutes instruction focusing on metalinguistic awareness can be very effective. Metalinguistic awareness based instruction may be effective for two reasons. First, enhancing metalinguistic awareness seemed to help learners to understand the importance of English prosody and provides learners a motivation to use prosody in their daily lives. Second, raising learners' metalinguistic awareness helps learners to pay more attention to sentence stress, which may promote learners' "noticing" (Schmidt, 1990) in language learning.

This study finds that the pitch height of the stressed constituents of the trans-MAET group has a statistically significant increase whereas the increase of the mono-MAET group was not statistically significant, suggesting that trans-MAET is more effective than mono-MAET. This result has two implications. First, this result reveals a limitation of mono-MAET: because of the

dynamic nature of prosody, learners could not memorize the pitch contour of each word and apply it to every sentence. Prompting the subjects to transfer similar prosodic functions/usages from their L1s can not only help them to understand the importance of prosody in English but also pick up the usage of English prosody by transferring prosodic usage from their L1s. Second, this result supports the translanguaging approach in prosody teaching: enhancing L2 English speakers' awareness of the similarities between Mandarin and English sentence stress may allow them to elicit prosody usage from their repertoire of languaging practices and apply to English.

This study shows that even though the pitch contours of Mandarin and English stressed constituents are different, both languages use increased pitch height to signal sentence stress. This result requires researchers and teachers to address the complexity of prosody and avoid overgeneralization of certain prosodic features. This result also suggests that more cross-linguistic research studies analyzing prosodic features from different perspectives should be conducted.

The results of this study also require teachers to take learners' L1s prosody into consideration when teaching English prosody. However, although some scholars discussed positive and negative transfer across languages, information regarding how each prosodic feature is used in different languages is not easily accessible to teachers. Teaching materials concluding cross-linguistic research findings are needed.

This study has certain limitations. First, the study had a limited number of participants. Follow-up studies should expand the number of participants to confirm the findings of this study. Furthermore, this study only investigated the pitch height of the stress words. Follow-up studies should investigate other prosodic features of the stressed and unstressed constituents including pitch range, duration, and intensity. Finally, this study used a read-aloud task to elicit data. Future study should investigate the efficacy of the proposed teaching methods using tasks eliciting spontaneous speech.

Overall, in investigating the efficacy of two pedagogy for prosody teaching, this study offers a possible innovation in teaching approach to pronunciation teachers and sheds light on possible future research avenue.

ABOUT THE AUTHOR

Di Liu is doctoral fellow at Boston University. He earned his MA degree on TESOL from New York University and is working with Dr. Marnie Reed on applied phonology and Computer-Assisted Language Learning. He has taught English in various contexts and has been working on several research projects including (1) technology-enhanced pronunciation research and teaching, (2) cross-linguistic investigation of Mandarin and English prosody, (3) constraints on intonation as a function of tone inventory structure, and (4) neurocognitive investigation of prosodic features. He has presented his studies at various conferences including TESOL, AAAL, WATESOL and PSLLT conference.

19043 Steeple Pl.

Germantown, MD, 20874

diliu@bu.edu

917-868-5796

REFERENCES

Anderson-Hsieh, J., Johnson, R., & Koehler, K., (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning 42,* 529–555.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341-345.

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

Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, *36*(4), 724-746.

Clennell, C. (1997). Raising the pedagogic status of discourse intonation teaching. *ELT Journal, 51*(2), 117-25.

Cruz-Ferreira, M. (1987). Non-native interpretive strategies for intonational meaning: An experimental study. *Sound Patterns in Second Language Acquisition*, 103-120.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *48*(3), 393-410.

García, O., and Li, W. (2014). *Translanguaging: Language, bilingualism and education*. New York: Palgrave Macmillan.

Gilbert, J. (2014). Myth 4: Intonation is Hard to Teach In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching*. (pp. 107-136). Ann Arbor, MI: University of Michigan Press:

Hahn, L. D. (2004). Primary stress and intelligibility: research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2)*,* 201-223.

Ip, M. H. K., & Cutler, A. (2016) Cross-language data on five types of prosodic focus. Proc. Speech Prosody 2016, 330-334. doi: 10.21437/SpeechProsody.2016.

Li, W. (2011) Moment analysis and translanguaging space: Discursive construction of identities by multilingual Chinese youth in Britain. *Multilingual Structures and Agencies, 43* (5), 1222–1235.

Mazak, C. M. (2017). Introduction: Theorizing Translanguaging Practices in Higher Education. In C. M. Mazak & K. S. Carroll (Eds.), *Translanguaging in higher education: Beyond monolingual ideologies.* (pp. 1-10). Bristol; Blue Ridge Summit, PA: Multilingual Matters.

Ouyang, I. C., and Kaiser, E. (2015). Prosody and information structure in a tone language: An investigation of Mandarin Chinese. *Language, Cognition and Neuroscience*, *30*(1-2), 57-72.

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

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reed, M., & Michaud, C. (2005). *Sound concepts: An integrated pronunciation course*. New York: McGraw-Hill.

Reed, M., & Michaud, C. (2015). Intonation in research and practice: The importance of metacognition. In M., Reed & J. Levis (Eds.), *The handbook of English pronunciation*. (pp. 454-470). West Sussex: UK: John Wiley & Sons.

Schmidt, R. W. (1990). The role of consciousness in second language learning1. *Applied linguistics*, *11*(2), 129-158.

Trofimovich, P., Kennedy S., & Foote, J. (2015). Variables affecting L2 pronunciation development. In M. Reed & J. Levis (Eds.) *The handbook of English pronunciation* (pp. 353-373), West Sussex, UK: John Wiley & Sons, Inc.

Wennerstrom, A. (1998). Intonation as cohesion in academic discourse: A study of Chinese speakers of English. *Studies in Second Language Acquisition, 20* (1), 1-25.

Williams, C. (1996) Secondary education: Teaching in the bilingual situation. In C. Williams, G. Lewis and C. Baker (Eds.), *The Language Policy: Taking Stock* (pp. 193–211). Llangefni: CAI.

APPENDIX: Read Aloud PASSAGE

I will start by defining the topic for today, which is individualism and collectivism. Individualism concerns the placing of personal goals ahead of group goals. And collectivism concerns placing group goals ahead of personal goals. So let's suppose you have a conflict at work about break time. Let's say your co-workers want longer breaks, but you want shorter breaks. If you're a collectivist, you'll give in to the group. But if you're an individualist, you'll go against the group.

First of all, there are many determinants of individualism and collectivism. Culture is a determinant, but it's only one of the determinants. But let me start with culture. Basically, the European cultures, particularly those in northwestern Europe, are highly individualistic. England is in northwest Europe, and it's typical of the individualist pattern. And the East Asian cultures, such as China and Japan, are much more typical of the collectivist pattern. But in between, you have different combinations of the patterns. And I'll discuss that in a minute. But let me mention some other determinants of individualism and collectivism.

One determinant that's very important is social class. There's a tendency for the upper classes to be more individualistic than the lower classes. In other words, people at the top of a social structure are more likely to think and behave like an individualist than those near the bottom of the structure. For example, if you look at the history of China, the emperor of China is more individualistic than the working class in China. So certain classes in a culture may be more individualistic than the entire population in a culture.

Zárate-Sández, G. (2018). "Was that a question?" Perception of utterance-final intonation among L2 learners of Spanish. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 76-86). Ames, IA: Iowa State University.

"WAS THAT A QUESTION?" PERCEPTION OF UTTERANCE-FINAL INTONATION AMONG L2 LEARNERS OF SPANISH

Germán Zárate-Sández, Western Michigan University

This study examined the perception of final boundary tones among 55 English-speaking learners of Spanish at three proficiency levels and compared them with Spanish-English early bilinguals, Spanish monolinguals, and English monolinguals. Perception was tested using an imitation task, aimed at capturing categorical perception effects. The stimuli consisted of a resynthesized utterance where the final boundary tone was vertically displaced ten times in 10-Hz increments. Participants listened to and imitated each stimulus twice while being recorded. Each final boundary tone was manually marked and extracted for all utterances participants produced. A one-way ANOVA was run for each group. Significant differences in the perception of the ten stimuli were found in all groups. Post-hoc analyses showed that (1) no bimodal categorical perception emerged; (2) stimuli clustered from falling (declarative) to rising tones (questions); (3) overall, all participants perceived final boundary as Spanish monolinguals did; (4) a third pattern suggestive of a suspended tone emerged for bilingual speakers and learners with very high proficiency; and (5) very-high proficiency learners' perception resembled that of bilingual speakers. Results confirm the robustness and perceptual salience of utterance-final pitch—which aids perception from early stages of acquisition—as well as a positive role for proficiency in the perception of the target form.

INTRODUCTION

The field of second language (L2) pronunciation has paid little attention to how learners perceive pitch in the L2. Such research is essential to inform models of L2 pronunciation such as Flege's (1995) Speech Learning Model (SLM), where attunement in L2 perception is a prerequisite for development of L2 production. Even though some studies have examined the perception of prosody as it relates to lexical tones (e.g., Broselow, Hurtig, & Ringen, 1987), few have looked at how intonation in the L2 is perceived. In an attempt to further our knowledge in this area, this study investigated how English-speaking learners perceive final boundary tones in unmarked, non-emphatic declarative utterances in Spanish, an utterance type believed to pose some challenges for English speakers learning Spanish (e.g., Nibert, 2005, 2006). While speakers of both languages normally use falling intonation in these utterances, some varieties of English (including General American English) produce a rising contour—usually referred to as a high rising terminal (HRT)—where a falling one would be expected. This process of *uptalk* has been extensively studied in English, especially from dialectal and sociolinguistic perspectives (see Warren, 2016, for a comprehensive description of this topic).

To my knowledge, Nibert (2005, 2006) is the only study that has addressed the perception of intonation in Spanish as an L2. English-speaking learners of Spanish at three proficiency levels

were tested on the effect of phrase accents (H- and L-) in Spanish to disambiguate utterances like [[*lilas*]$_{H-}$ [*y lirios amarillos*]$_{L-}$] 'lilacs and yellow irises', [[*lilas y lirios*]$_{H-}$ [*amarillos*]$_{L-}$] 'yellow lilacs and yellow irises', and [[*lilas y lirios amarillos*]$_{L-}$] where the lack of medial H- allows for either interpretation. Results showed a clear positive effect of proficiency level on the perception of this particular tone. The data suggested that there can be a "gradual development or restructuring of L2 interlanguage grammar toward a more restrictive and native-like state" (2006, p. 146).

Though they examined L2s other than Spanish, two important studies investigated how final boundary tones are perceived by L2 learners. In Cruz–Ferreira (1987), 30 English speakers learning Portuguese and 30 Portuguese speakers learning English heard 60 pairs of sentences that differed in their intonational pattern (e.g., *Didn't John en`joy it* vs. *Didn't John en¸joy it*) and were asked if the sentences were the same or different while also matching the sentence to glosses. Cruz–Ferreira found that learners used at least three interpretive strategies when resolving the task. The *transfer* strategy was used when the L1 and L2 had similar intonation structures and hence the L1 meaning was assigned to the L2. The *pitch height* strategy stated that L2 learners identified L2 tones correctly when they broadly belonged to a certain category (for example, rise vs. fall). Finally, the *lexico-syntactic* strategy occurred when learners assigned the less marked interpretation to the sentences based only on their lexical or grammatical pattern. In another study on perception, Grabe, Rosner, García–Albea, and Zhou (2003) conducted two experiments aimed at testing the perception of falling and rising final intonation in English among adult speakers of Spanish and Chinese. Participants heard a pair of stimuli (the phrase *Melanie Maloney*) that differed only in the final pitch movement (rising or falling) and rated the stimuli as same or different. The second experiment was similar but used only contour movements without any speech. As predicted, all three groups made a clear distinction between falling and rising contours in both speech and non-speech stimuli. The authors attributed this to a purported universal (that is, non-language specific) auditory mechanism used in the perception of pitch (Bolinger, 1978; Cruttenden, 1981).

In view of the different intonational pattern in declarative utterances in English and Spanish, and given our limited knowledge on how utterance-final intonation in Spanish is perceived by L2 learners, the following research questions were posed:

(1) How do English-speaking learners of Spanish at different proficiency levels perceive the final boundary tone of Spanish declarative utterances?
(2) How does their performance compare to that of monolingual native speakers of Spanish, monolingual native speakers of English, and early Spanish-English bilinguals?

METHODOLOGY

Participants

Initial recruitment was conducted from third-semester and sixth semester Spanish classes at two universities in the U.S. In addition, students pursuing a Master's or Ph.D. in Spanish at one of the universities were also recruited. After initial recruitment, participants' proficiency was measured with the Spanish Elicited Imitation Task (EIT) test (Ortega, 2000), which is argued to be a reliable and valid tool to measure L2 oral proficiency (see Bowden, 2016, for discussion). For the sake of space, details about administration and scoring are omitted here, but they followed

the same procedures described in Bowden (2016) and Ortega (2000). Final proficiency was hence operationalized based on two criteria: previous experience and scores on the EIT. From a possible range of scores between 0 and 120, the means for the three groups were the following: 57.96 (*SD* = 12.24) for the third-semester group, 93.33 (*SD* = 11.65) for the sixth-semester group, and 116.40 (*SD* = 3.78) for the group of graduate students. In order to maintain homogeneity in each group, outliers (two standard deviations above or below the group mean) were eliminated from the study. For this and other exclusion criteria (e.g., not completing all parts of the study or having extended stays in Spanish-speaking countries), 15 participants were excluded from the original pool, resulting in a total of 55 learners grouped into three levels: intermediate proficiency (IP, *n* = 17), comprised of third-semester Spanish students with EIT scores in the 42–82 range, high proficiency (HP, *n* = 20), comprised of sixth-semester students with EIT scores in the 83–107 range, and very high proficiency (VHP, *n* = 18), comprised of graduate students with EIT scores in the 109–120 range.

In addition, L2 speakers were compared with three groups: Spanish native speakers (SNS, *n* = 17), English (monolingual) native speakers (ENS, *n* = 17), and English-Spanish bilingual speakers (BS, *n* = 16). SNS speakers were from various Spanish dialects in order to represent the variety of dialects to which learners in this study had likely been exposed: Southern Cone (*n* = 4), Andean (*n* = 4), Mexican (*n* = 3), Central American (*n* = 3), and Peninsular (*n* = 3). ENS participants were recruited from undergraduate classes at the same university. They were all monolingual native speakers of American English. The BS group was comprised of heritage speakers, that is, early bilinguals who were raised in families that speak a minority language. The inclusion of this group attempted to expand the comparison beyond typical monolingual norms (Ortega, 2009) and reflect the bilingual nature of interlanguage development. In pronunciation, recent studies have also suggested that bilingual speakers are appropriate comparison subjects for L2 learners (e.g., Sakai, 2018).

Perception task

Perception of final boundary tone was examined through an imitation task, which has been shown to be an effective method to investigate categorical perception (CP) effects in intonation (Dilley & Brown, 2007; Gussenhoven, 1999, 2004). In this task, participants usually listen to one stimulus at a time from a continuum and are asked to reproduce it out loud. If there is a categorical distinction, participants will not reproduce the entire continuum they hear and utterances tend to group in a bimodal distribution. The underlying principle and possible outcomes resemble those of discrimination and identification tasks, but the imitation task offers the advantage that no preexisting categories are required (as in an identification task) and only one stimulus is heard at a time (compared to at least two, as in a discrimination task), thus posing fewer challenges for memory retention (Gussenhoven, 2006). For intonation, Pierrehumbert and Steele (1989) were among the first to use this task and it has been employed to investigate pitch alignment (e.g., Pierrehumbert & Steele, 1989; Redi, 2003) and pitch height (e.g., Dilley, 2005; Dilley & Brown, 2007).

The declarative sentence *La nena lloraba* 'The girl was crying' was used as the source utterance to create the stimuli. This sentence was selected for its simple lexical composition, thus posing few challenges for L2 speakers. Additionally, the following phonotactic and segmental features were taken into consideration when choosing the sentence: (1) all consonants are voiced, which is an important requirement to obtain an uninterrupted pitch track, (2) only mid and low vowels (that is, [e, o, a]) are used (high vowels may produce undesired pitch alterations by raising the

larynx, Gussenhoven, 2004), (3) syllables have a CV structure (no consonant clusters and no codas) in order to minimize potential difficulties for L2 speakers while also maximizing sonority necessary for pitch detection, and (4) content words are stressed on the penultimate syllable, which is the unmarked stress pattern in Spanish for words ending in a vowel.

The sentence was recorded by a native speaker of western Argentinian Spanish using a high-sensitivity microphone attached to a personal computer and the acoustic software Praat (version 5.1.25) at a sampling rate of 44 kHz. The file was resynthesized using the *pitch-synchronous overlap-add* (PSOLA) method included in Praat. This technique allows for stylization of the pitch track by reducing it to critical pitch points while keeping the overall curve shape. The resulting contour was inspected acoustically and auditorily before manipulation in order to check it was naturally-sounding. Relevant pitch points were reduced to the initial (InP) and final (FiP) points of the last intonational move. Height (117 Hz) and timing (middle point of [a] in syllable *ra*) of InP were kept as in source utterance. FiP was anchored at the last regular glottal pulse detected in the spectrogram, therefore excluding visible and audible creaky phonation at the end of the utterance. Finally, FiP was moved upwards seven times and downwards twice, in 10-Hz increments, creating 10 stimuli spanning over 90 Hz (See Figure 1), downsampled to 22 kHz and saved individually as WAV sound files. Two native speakers of Spanish listened to the resulting stimuli and judged them as naturally-sounding (that is, potentially produced by a human as opposed to digitally created).

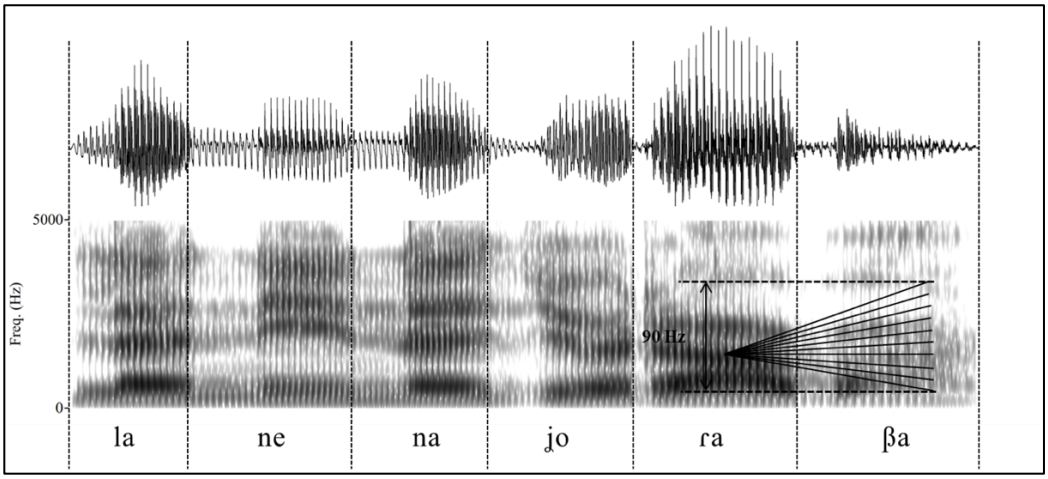


Figure 1. Schematic representation of resynthesized stimuli used in perception task.

Procedure

Participants were presented with the ten stimuli twice, in two randomized blocks, resulting in 20 repetitions per participant. Participants were instructed to reproduce the utterance they heard as faithfully as possible. They were asked to focus on the pronunciation of the sentence and encouraged to imitate it within a comfortable pitch range. Participants first listened to a block of five practice utterances, which they could repeat until they felt comfortable with the task. For the trial blocks, participants had the possibility of saying the utterance again if they hesitated, paused, or considered that their output was not faithful to the stimulus they had heard. The stimuli were delivered over high-fidelity headphones and presented through E-prime software.

Recordings were made in a sound-proof room, with a high-sensitivity, head-mounted microphone attached to a personal computer and using Audacity (version 1.3.14), at a sampling rate of 32 kHz.

In order to collect information about the meaning participants assigned to the stimuli, upon completion of the task the researcher asked participants the following questions: (1) *Did the sentences sound the same or different to you?* (2) *If different, how so?* and (3) *In what contexts would you use the sentence(s) you heard?* During this interview, answers were audio-recorded and participants were debriefed on the nature of the stimuli and the goals of the study. Finally, participants completed a questionnaire aimed at collecting demographic information and data on previous language experience.

Analysis

After auditory and visual inspection of spectrograms and intonation curves of 2,100 utterances (20 utterances x 105 subjects), 555 (26.43% of the total) were removed from the analysis due to excessive creaky voice, hesitations, and uncommonly flat global pitch contours. Final creaky voicing, especially among male speakers, was the main reason for elimination. For this reason, two male speakers were removed altogether as more than 50% of their utterances contained creaky voice. In total, 1,545 utterances were analyzed (249 for IP group, 300 for HP, 276 for VHP, 252 for SNS, 231 for ENS, and 237 for BS). For these utterances, the final boundary tone was defined as the last non-spurious f_0 point in the pitch track generated by Praat (see Arvaniti & Ladd, 2009; Henriksen, 2014, for similar procedures). Extracted f_0 values were converted to equivalent rectangular bandwidth (ERB) units in order to normalize for differences among speakers, in particular between males and females. This normalization procedure describes more accurately the relationship between f_0 and perceived pitch and has been used increasingly in intonation research (e.g., Arvaniti & Ladd, 2009; Colantoni & Gurlekian, 2004; Simonet, 2009).

RESULTS

Boxplots (Figure 2) for each group present the descriptive statistics of the imitation task. The X axis represents the ten points in the stimuli, while the Y axis represents the normalized f_0 scaling of the final tone, expressed in ERB units.

In general terms, boxplots for perception of final boundary tones reveal that participants perceived increments in final f_0 in a gradient manner. That is, descriptive statistics do not suggest clear CP differences in the perception of boundary tone stimuli. A one-way ANOVA was run for each group in order to explore further potential CP effects in the data (Table 1). The 10 points in the stimuli served as the independent variable, and ERB units served as the dependent variable.

Results of the omnibus one-way ANOVA yielded significant difference in all six cases. Notice, however, the reduced effect sizes reported in η^2. This could be attributed to the rather small sample due to the elimination of 555 data points. Nonetheless, since all tests found significant differences between stimuli, a post-hoc analysis (Tukey's) was run in order to determine the precise location of these differences. Results are presented graphically in Figure 3.

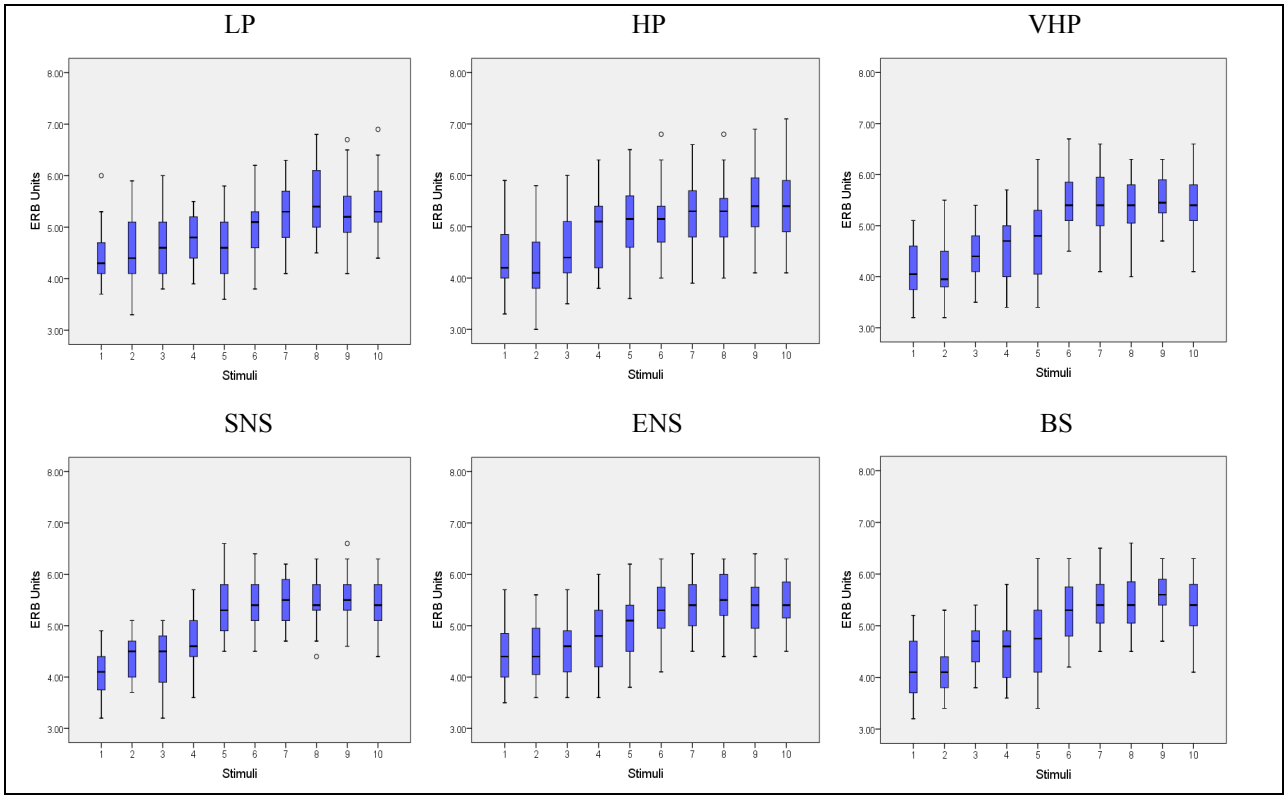


Figure 2. Boxplots with results in imitation task for all groups.

Table 1

One-way ANOVA for Results in Final Boundary Tone for All Groups

Group	*F*	*df*	Sig.	Partial η^2	Observed power
LP	11.09	9 - 239	.000	.29	1.00
HP	12.48	9 - 290	.000	.28	1.00
VHP	27.81	9 - 266	.000	.48	1.00
SNS	30.85	9 - 242	.000	.53	1.00
ENS	13.65	9 - 221	.000	.36	1.00
BS	23.34	9 - 227	.000	.48	1.00

The results of post-hoc comparisons revealed that no group achieved CP effects in the perception of the stimuli as traditionally defined, since all groups had more than two clusters and there was some degree of overlap among these clusters. Interestingly, however, if we are willing to

consider the first two overlapping groups (blue and red in Figure 3) as one, the VHP, SNS, and BS groups did have one clear cut off point between stimuli 4 and 5 for SNS, and between stimuli 5 and 6 for VHP learners and BS. Otherwise, the stimuli for all groups were arranged from falling (non-emphatic declarative) to rising (questions), which corresponded to the declarative versus question interpretations, respectively, that the vast majority of participants reported in the debriefing interview.

Group	1	2	3	4	5	6	7	8	9	10
	Stimuli for final boundary tone									
LP	4.41	4.54	4.58	4.78	4.63	4.98	5.24	5.47	5.33	5.40
HP	4.40	4.24	4.48	4.95	5. 14	5.17	5.24	5.28	5.42	5.41
VHP	4.14	4.13	4.43	4. 55	4.74	5.56	5.48	5.39	5.54	5.49
SNS	4.11	4.38	4.36	4.74	5.40	5.43	5.51	5.52	5.56	5.47
ENS	4.44	4.48	4.54	4.78	5.02	5.31	5. 40	5.51	5.38	5.48
BS	4.14	4.20	4.63	4.54	4.74	5.27	5.50	5.50	5.60	5.37

Figure 3. Results for post-hoc comparisons in final boundary tone stimuli.

Note. Scores represent mean ERB units produced for each stimulus, by group. Scores grouped in the same color indicate that means are statistically equal (Tukey's test).

DISCUSSION AND CONCLUSION

The results for perception of boundary tone showed that all participants were able to perceive broad increments in pitch in an overall similar manner. The data from debriefing interviews also determined that that these differences in pitch have at least two interpretations: questions (for rising tones) or declaratives (for falling tones). Indeed, regardless of proficiency, participants perceived overall pitch height and excursions accurately, a perception process that Cruz–Ferreira (1987) noticed in L2 speakers and called *pitch height strategy*. In her opinion, learners "seem to be sensitive to the gross phonetic shape of the pattern" (p. 112), which is supported by scholars such as Cruttenden (1981) who assign the general universal meanings of *open statement* to rising contours and *closed statement* to falling ones. This rudimentary strategy may constitute the first phase in perceiving and shaping the L2 intonational system. This finding seems to contradict the initial hypothesis that this intonational pattern in Spanish would pose a challenge for English speakers. The current study appears to demonstrate that perception of Spanish boundary tone might be an aspect on interlanguage phonology that naturally develops from an early stage and approaches native-like processing at high proficiency levels. This can be achieved in an

instructed setting and with minimal to no explicit pronunciation instruction, as was the case for participants in this study.[1]

The finding that even monolingual English speakers processed the stimuli in a broadly similar manner as other groups provides further evidence for the robustness of the perceptual strategy under investigation: it can be activated successfully even if the hearer does not possess familiarity with the L2. As reported in the background questionnaire, these speakers had negligible contact with spoken Spanish in their daily lives and had received no instruction in Spanish during high school. Yet, they perceived global falls and rises similarly to other participants in this study. It could be argued that they were processing the auditory stimuli not as Spanish intonation but merely as changes in pitch excursions. These results are in line with Grabe et al. (2003), where perception of final intonation was broadly the same for both speech and non-speech stimuli. The fact that the broad distinction between falls and rises in English is the same as in Spanish—and in many if not most other languages, as argued by Cruttenden (1981)—could have also aided monolingual participants in their perception of Spanish intonation.

Upon closer examination, however, there were differences between monolingual English speakers and the other groups. The results of the debriefing interview revealed that *question* and *statement* were the only two possible interpretations that intermediate proficiency, high proficiency, and English speakers assigned to boundary tone stimuli. However, a subset of participants in the groups of very high proficiency L2 speakers, bilingual speakers, and Spanish native speakers reported that certain utterances seemed to be 'unfinished' or 'as if the speaker wanted to say something else.' We could hypothesize that this third interpretation may be attributed to a level final tone, pragmatically interpreted as a suspension tone. This possibility appears to be corroborated in the perception data: these three groups actually perceived three distinct groups of stimuli (see Figure 3), where the middle group—represented in red and spanning stimuli 2–5—could be linked to this purported suspension tone. This is an appealing possibility and would suggest that, as proficiency in Spanish increases, perception of final tone is attuned to three basic utterance types: declarative, suspended, and interrogative. Authors such as Levis (1999) have also noticed the role of this third utterance-final pattern in interlanguage phonology. In regards to the English falling-rising contour he observed, Levis stated that because "one function of intonation in conversation is to communicate whether a speaker has finished a turn, falling-rising intonation is a key communicative resource for speakers" (p. 43).

It is also worth noticing that only the SNS, VHP, and BS groups—the groups with the highest proficiency in Spanish—achieved some level of categorical perception, even if not produced with the traditional bimodal distribution. More interestingly, the pattern of clustering for the VHP group was exactly the same as that of BS. These findings contribute to research that suggests high proficiency in a foreign language may result in native-like processing (see Bowden, Steinhauer, Sanz, & Ullman, 2013, for an example of syntactic processing). To my knowledge, this current study is the first to obtain such results for perception of L2 prosody.

[1] A set of questions in the background questionnaire addressed the type and amount of explicit pronunciation instruction participants had received. Even though most said they had received explicit explanations for Spanish segments known to be difficult for English speakers, none recalled ever being taught about or receiving feedback on aspects of Spanish intonation.

Finally, we need to consider the context where L2 speakers in this study learned Spanish. All participants started learning Spanish as teenagers in instructed settings. This type of instruction continued into college, where only some participants in the very-high proficiency group also had the opportunity of extended stays in a Spanish-speaking country. Given this general profile, it is safe to assume that the L2 input participants had received was not only limited in amount but also restricted to a combination of native and non-native sources (classmates and non-native instructors, for example). The background questionnaire revealed that in fact all L2 participants received substantial input from non-native speakers of Spanish yet showed significant development in their perception of L2 intonation. In other words, it appears to be possible for college-level students of Spanish to improve their perception of Spanish intonation *despite* limited access to native L2 input. These results underscore the beneficial role of instructed contexts in the development of certain aspects of L2 intonation.

LIMITATIONS

The imitation task employed in this study is a methodological compromise which, as discussed above, has been deemed an effective tool to examine perception of intonation, but the fact that it tests *perception* while relying on *production* remains problematic. Even though the selected utterance was simple and did not pose challenges for the speakers, all of whom had at least three semesters of instruction, some participants were simply poor imitators, as also noticed by Dilley (2010), and felt uncomfortable with the task. The elimination of 26% of the data points from the analysis may reflect participants' insecurity while performing the task. Future research should build upon findings in this study and triangulate results from an imitation task with data from different perception tasks. Additionally, the stimuli can be diversified to include, for instance, utterances of varying length and stress patterns.

ABOUT THE AUTHOR

Germán Zárate–Sández (Ph.D. Georgetown University, 2015) is assistant professor of applied linguistics, teacher trainer, and coordinator of basic and intermediate Spanish instruction at Western Michigan University. His current research examines the acquisition and use of intonation in Spanish as a second language with an emphasis on the link between perception and production of prosody.

1210 Arcadia Loop
Sprau Tower - Department of Spanish
Western Michigan University
Kalamazoo MI, 49008-5338
USA
german.zarate-sandez@wmich.edu

REFERENCES

Arvaniti, A., & Ladd, D. R. (2009). Greek wh-questions and the phonology of intonation. *Phonology, 26*(1), 43–74.

Bolinger, D. (1978). Intonation across languages. In J. P. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language. Volume 2: Phonology* (pp. 471–524). Stanford: Stanford University Press.

Bowden, H. (2016). Assessing second-language oral proficiency for research: The Spanish Elicited Imitation Task. *Studies in Second Language Acquisition, 38*(4), 647–675.

Bowden, H. W., Steinhauer, K., Sanz, C., & Ullman, M. T. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia, 51*(13), 2492–2511.

Broselow, E., Hurtig, R. R., & Ringen, C. (1987). The perception of second language prosody. In G. Ioup, & S. H. Weinberger (Eds.), *Interlanguage phonology: The acquisition of a second language sound system* (pp. 350–361). Cambridge, MA: Newbury House.

Colantoni, L., & Gurlekian, J. (2004). Convergence and intonation: Historical evidence from Buenos Aires. *Bilingualism: Language and Cognition, 7*(2), 107–119.

Cruttenden, A. (1981). Falls and rises: Meanings and universals. *Journal of Linguistics, 17*(1), 77–91.

Cruz–Ferreira, M. (1987). Non-native strategies for intonational meaning: An experimental study. In A. James & J. Leather, J. (Eds.), *Sound patterns in second language acquisition* (pp. 103–120). Dordrecht: Foris.

Dilley, L. C. (2005). *The phonetics and phonology of tonal systems* (Doctoral dissertation). Massachusetts Institute of Technology.

Dilley, L. C. (2010). Pitch range variation in English tonal contrasts: Continuous or categorical? *Phonetica, 67,* 63–81.

Dilley, L. C., & Brown, M. (2007). Effects of pitch range variation on f0 extrema in an imitation task. *Journal of Phonetics, 35,* 523–551.

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

Grabe, E., Roser, B. S., García–Albea, J. E., & Zhou, X. (2003). Perception of English intonation by English, Spanish, and Chinese listeners. *Language and Speech, 46,* 375–401.

Gussenhoven, C. (1999). Discreteness and gradience in intonational contrasts. *Language & Speech, 42*(2), 283–305.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. New York: Cambridge University Press.

Gussenhoven, C. (2006). Experimental approaches to establishing discreteness in intonational contrasts. In S. Sudhoff et al. (Eds.), *Methods in empirical prosody research* (pp. 321–334). Berlin: Walter de Gruyter.

Henriksen, N. (2014). Initial peaks and final falls in the intonation of Manchego Spanish wh-questions. *Probus 26*(1), 83–133.

Levis, J. M. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly, 33*(1), 37–63.

Nibert, H. (2005). The acquisition of the phrase accent by intermediate and advanced adult learners of Spanish as a second language. In D. Eddington (Ed.), *Selected Proceedings of the 6th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages* (pp. 108–122). Somerville, MA: Cascadilla Proceedings Project.

Nibert, H. (2006). The acquisition of the phrase accent by beginning adult learners of Spanish as a second language. In M. Díaz–Campos (Ed.), *Selected Proceedings of the 2nd Laboratory Approaches to Spanish Phonetics and Phonology* (pp. 131–148) Somerville, MA: Cascadilla Proceedings Project.

Ortega, L. (2000). *Understanding syntactic complexity: The measurement of change in the syntax of instructed L2 Spanish learners* (Doctoral dissertation). University of Hawaii at Manoa.

Ortega, L. (2009). *Understanding second language acquisition*. London: Routledge.

Pierrehumbert, J., & Steele, S. (1989). Categories of tonal alignment in English. *Phonetica, 46,* 181–196.

Redi, L. (2003). Categorical effects in production of pitch contours in English. In *Proceedings of the 15th International Congress of the Phonetic Sciences* (pp. 2921–2924). Barcelona.

Sakai, M. (2018). Moving towards a bilingual baseline in second language phonetic research. *Journal of Second Language Pronunciation, 4*(1), 11–45.

Simonet, M. (2009) Nuclear pitch accents in Majorcan Catalan declaratives: Phonology, phonetics, diachrony. *Studies in Hispanic and Lusophone Linguistics, 2*, 77–115.

Warren, P. (2016). *Uptalk: The phenomenon of rising intonation.* Cambridge: Cambridge University Press.

Dao, D. (2018). Learners' perspectives on English pronunciation teaching and learning: A preliminary study in the Vietnamese context. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 86-99). Ames, IA: Iowa State University.

LEARNERS' PERSPECTIVES ON ENGLISH PRONUNCIATION TEACHING AND LEARNING: A PRELIMINARY STUDY IN THE VIETNAMESE CONTEXT

Duc Nguyen Anh Dao, University of Nottingham, Malaysia Campus

In a context where many L2 teachers follow their intuition in making decisions related to pronunciation teaching (Derwing & Munro, 2005; Levis, 2005), an insight into the learner's view of what is happening in the classroom might help teachers become better informed so as to make better choices. The study aims to find out what Vietnamese adult learners think about the current teaching and learning of English pronunciation. 38 learners participated in semi-structured interviews which centered around issues related to their learning targets, learning difficulties as well as their expectations for and evaluations of the pronunciation instruction available to them. Findings reveal that learners aimed at achieving a native like accent, which was considered a proof of success in learning. They also tended to have similar problems in acquiring the phonological features of English. Although almost all of them highly valued the role of instruction, many were not satisfied with the current teaching practices and called for more effective instruction with regards to focus, type of classroom activities and availability of feedback. The study's findings and implications are expected to help curriculum designers and teachers determine what to teach and how to teach it for better outcomes.

INTRODUCTION

English teaching in Vietnam has undergone great changes since September 30th, 2008 when the National Foreign Languages 2020 Project was approved by the Ministry of Education and Training. A budget of 9,400 billion VND would be spent on reforming language teaching and learning. Authorities declared that this project would help Vietnamese young citizens to gain the ability to use a foreign language independently and confidently for communication, work or study in a multilingual and multicultural environment ("National Foreign Languages 2020 Project," n.d.).

Since the launch of the project, English teaching has experienced massive reforms in all aspects (curriculum design, teacher training and re-training, textbook writing), at all levels and in all types of educational institution (Hoang, 2010). New policies derived from the project require Vietnamese teachers of English to prove that they are linguistically qualified for the job, otherwise they will have to be retrained, and Vietnamese students to demonstrate that they can use English satisfactorily in order to obtain their degrees. Standardized tests such as IELTS, FCE and CAE are commonly used for those two purposes. Because the international testing systems assess language users in four skills, including speaking and listening, pronunciation is no longer a "neglected aspect" (Alghazo, 2015) in English language teaching.

Disappointingly, such huge investments have not produced expected outcomes. At the end of 2016, Minister of Education and Training Phung Xuan Nha admitted that Project 2020 had not achieved its targets (Vo & Hoai, 2016). Long-lasting problems related to teacher quality and learning

outcomes remain unsolved. Specifically in the field of pronunciation, Vietnamese accented English is still considered unintelligible for native English speaking listeners and even for Vietnamese listeners (Cunningham, 2009, 2013). The current study aims to examine the current teaching of English pronunciation in Vietnam in the hope to understand why little improvement has been made after eight years under reformation.

Recently, greater research interest has been shown in L2 pronunciation teaching and learning; however, little has been done from the learner's perspective. Alghazo (2015) claims that "one dimension of pronunciation teaching which has received hardly any attention to date is that of learner cognitions or beliefs about the way teachers of L2 pronunciation should approach this sub-skill" (p.63). The present study aims to address this gap, looking at the current teaching practices from the point of view of the learners.

BACKGROUND

Learning goals

When the goals of English teaching and learning are taken into account, teachers, and probably learners have to make a choice between two paradigms: nativeness or intelligibility. The nativeness principle holds that learners should aim and are able to achieve native-like pronunciation (Levis, 2005). However, this approach to the selection of a model is intuitive rather than empirical, and can be based on sociocultural, political or market-driven choices (Setter, 2008). Therefore, it has been considered unrealistic, unnecessary, and undesirable to both teachers and learners (Ketabi & Saeb, 2015; Moghaddam, Nasiri, Zarea, & Sepehrinia, 2012; Murphy, 2014).

Since the 1990s, the emergence of English as an international language has commanded that the goals of pronunciation teaching and learning be reconsidered; therefore, the intelligibility principle has received more attention. This principle postulates that L2 speakers just need to be understandable, communication can be successful when their accents are noticeable or even strong, and there is no clear correlation between accent and understanding (Munro & Derwing, 1995).

The recommendation that intelligibility be the target of L2 pronunciation teaching and learning does not necessarily mean that a native-like accent should never be the goal of learners. First, what accent to aim at should be the choice of the individual learners, not the teachers, the school or the government (Setter, 2008). Second, there is nothing wrong if learners themselves aim to achieve native-like pronunciation, as this goal may lead to greater learning motivation. Third, learners' needs may differ depending on the contexts in which they communicate, so teaching goals should also vary accordingly (Rogerson-Revell, 2011). The current study attempts to identify what Vietnamese learners want to achieve in learning English pronunciation as this information will help teachers determine their approach to teaching.

Potential learning difficulties

Gilakjani & Ahmadi (2011) realize that many second language learners have major difficulties with pronunciation even after a long time of learning the language. Researchers have attempted to predict and analyse areas of difficulty by means of contrastive analysis or error analysis. Several studies have been conducted on common problems Vietnamese learners of English face in learning pronunciation. Ha (2005)'s error analysis reveals that Vietnamese learners often omit *l, ʤ, r, s, i, ei*, and *k* in medial positions and *z, s, t, v, ks*, and *ʤ* in final positions but tend to add *s* and *z* at the end of a syllable or a word. Moreover, they find it difficult to produce some English sounds and

thus replace them with existing Vietnamese sounds. Some typical examples of this phenomenon are the English *ʧ* and the Vietnamese *t* or *tr*, the English *ð* and *θ* and the corresponding Vietnamese *z* or *d* and *s* or *t*. Some other works specifically focus on how Vietnamese learners deal with final consonants and final consonant clusters (N. Nguyen, 2002; T. H. Nguyen, 2002; Osburne, 1996). In addition, Nguyen (1998) attempted to validate the claim that English sounds made by Vietnamese learners are too short, investigating learners' omission of final sounds and distinction of long and short vowels.

It is noticeable that researchers have analysed data obtained from their own knowledge of the L1 and from their observations of learners' performance; very few empirical studies have been conducted from learners' perspectives to identify the struggles they actually undergo. Only when we are informed by both the researchers / teachers and the learners about potential difficulties will we be able to plan for more effective learning.

Teaching English pronunciation in the classroom

Even when L2 teachers are well aware of the importance of integrating pronunciation into the language classroom, they still find it formidable to teach it as "there is no agreed upon system of deciding what to teach, and when and how to do it" (Darcy, Ewert, & Lidster, 2012). The integrative approach imposes high demands on the design of pronunciation syllabi and lessons, which may also intimidate the teachers. Celce-Murcia et al. (2010)'s propose a communicative framework in which attention is gradually shifted from an initial focus on form towards incorporating more meaning in a sequence of activities, with the provision of corrective feedback. In addition, in a communicative classroom, the teacher should work as a pronunciation coach who assists learners by supplying information, giving models from time to time, offering suggestions and constructive feedback about performance, and providing practice opportunities and encouragement (Morley, 1991).

In the new era of language teaching when pronunciation has become an integral part of not only the whole curriculum but also every single lesson, modern techniques and activities have been devised in the hope to bring more success in teaching and learning L2 pronunciation. However, Lear (2011) admits that "there is a significant disparity between learner and teacher beliefs about the use of language learning activities" (p.131). While a large body of research has been done from the point of view of teachers, learners have rarely been asked for their opinions about what activities should be used or what they find useful. The present study, hence, will be focusing on this issue.

THE PRESENT STUDY

Research questions

1. What do Vietnamese learners aim to achieve in learning English pronunciation?
2. What difficulties do they encounter in learning pronunciation?
3. What are their perceptions of the current pronunciation teaching?

Participants

38 adult learners aged 17 and above were recruited on a voluntary basis from two institutions: Banking University of Ho Chi Minh City (BUH) and Vietnam-USA English Centers (VUS). Those from BUH all majored in English while their counterparts from VUS were more diverse. The

participants were categorized into three groups (BUH1, BUH2, and VUS) with background information summarized in Table 1.

Table 1

Participants' background

Group	N	Educational level			English proficiency	Amount of previous pronunciation instruction
		1st year	*2nd year*	*High school*		
BUH1	18	18	0	0	Intermediate	None
BUH2	10	0	10	0	Upper-Intermediate	1 course of 45 hours
VUS	10	2	4	4	Intermediate	Varied

Procedure

Invitations to join the study were sent to BUH students and those interested contacted me to set appointments. On the other hand, VUS learners were approached individually before their class meetings or during break times. The interviews, which were conducted in Vietnamese, lasted from 4 to 10 minutes, and were recorded with the participants' approvals. The interviews were semi-structured with a couple of questions dedicated to each research problem. In addressing the last problem regarding learners' perceptions of the current pronunciation teaching, one of the questions used was adjusted to match the amount and type of previous instruction each group received. (For a complete record of the interview questions, see the Appendix.)

RESULTS

Learning goals

Unsurprisingly, when asked whether they would like to be mistaken as native speakers of English, a vast majority of participants (31 out of 38) gave a positive answer; only three responded negatively. Two learners from group BUH1 said they would say both Yes and No, one student from group BUH2 expressed indifference, and the participant from group VUS had no idea at all. The participants gave very similar explanations. Proof of great achievements, competence and hard work and pride in oneself for success in learning were the reasons why most learners would love to have native-like accents. In contrast, the remaining participants claimed that they did not feel right and just wanted to be themselves.

Regarding which accent of English should be considered the model for learning, 23 participants stated a preference for American English while 13 voted for British English. The same pattern was found across the three groups. Two learners, however, answered that it did not matter as to which model should be followed and that they preferred whatever accent that was easier to listen to. The results are illustrated in Figure 1.

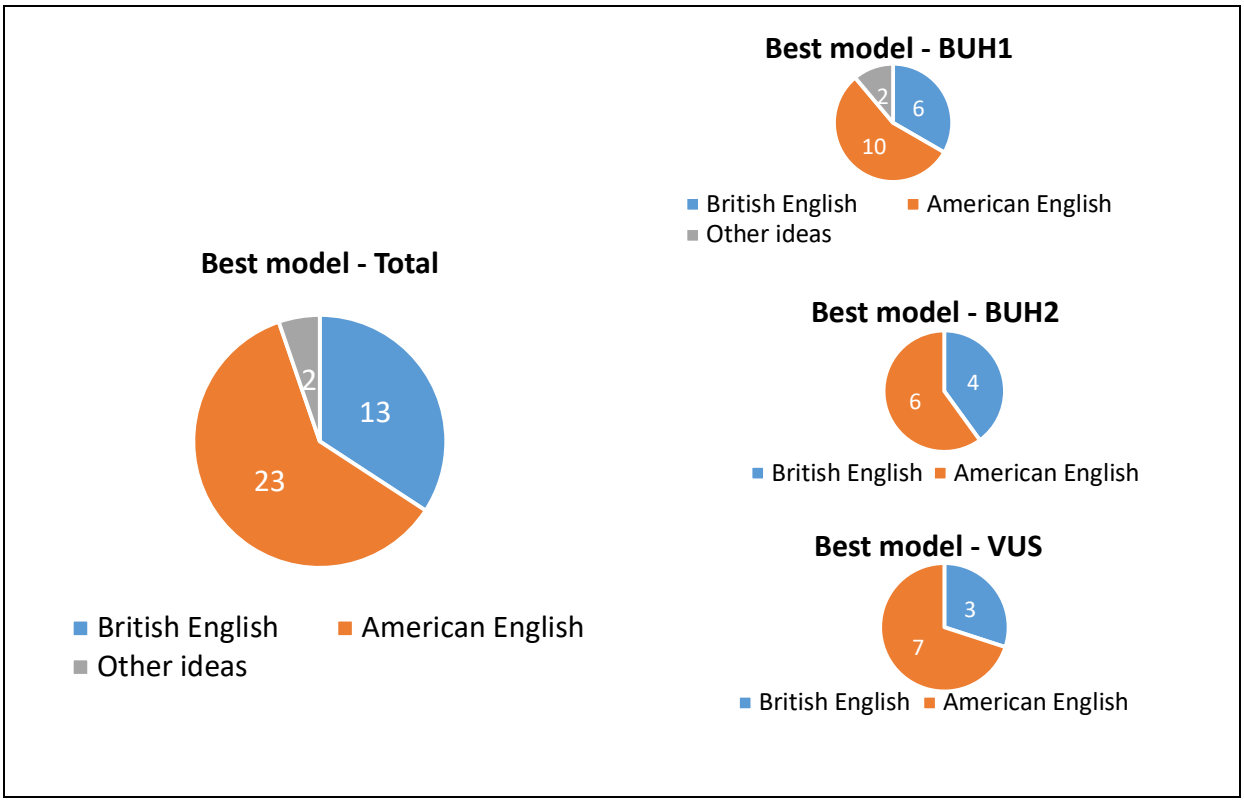


Figure 1. Best model of English

The participants provided a number of reasons for their choices, among which ease of listening and production was the most common. For those in favor of the British accent, two other important issues were also identified: it was the standard while other accents were not, and it was the origin of the English language. Interestingly, one student explained that this accent should be the model because it was more difficult and thus was worth learning. The details are shown in Figure 2.

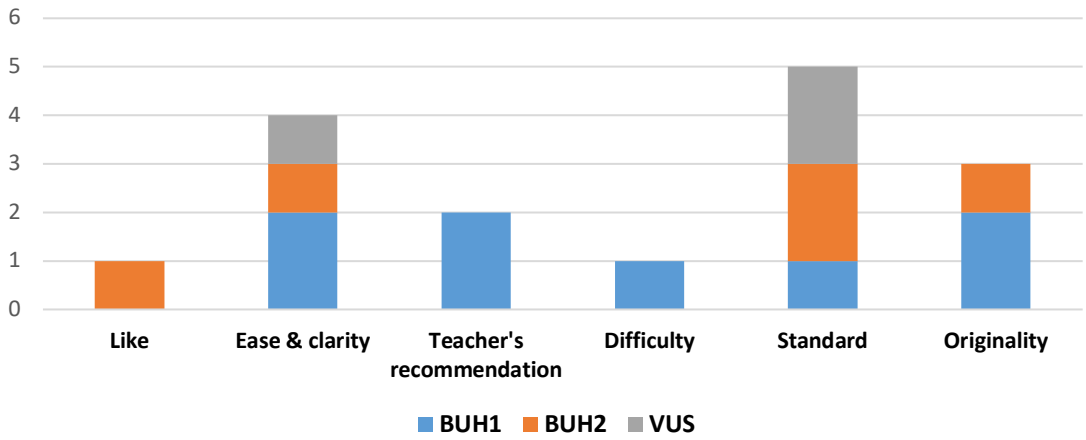


Figure 2. Reasons for choosing British English

As regard the preference for the American accent, the majority of participants explained that it was easier to learn. The second most common reason was that it was more popularly used in communication, textbooks, and movies. Two participants gave quite personal opinions, being that

the accent sounded more beautiful, and that their family lived in the US. A complete record of the data is presented in Figure 3.

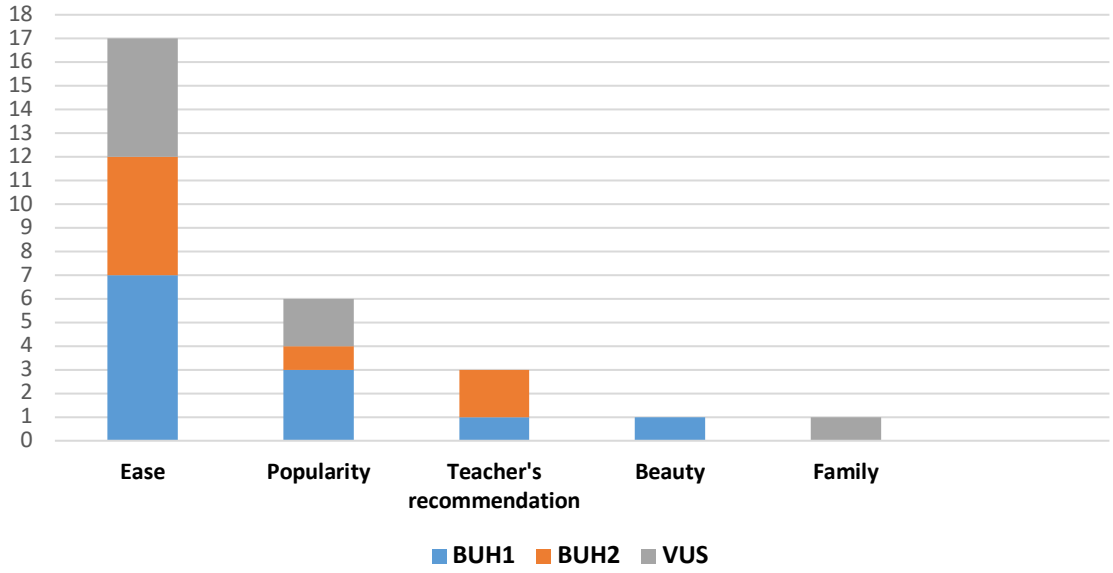


Figure 3. Reasons for choosing American English

Learning difficulties

The participants mentioned a variety of difficulties which were grouped into four categories: the first two consisted of issues pertaining to the types of phonological features (segmental vs suprasegmental), the third one included issues related to instruction and the last one comprised other problems. Table 2 lists the difficulties and the number of times they were mentioned by the participants. In general, more segmental problems were named than suprasegmental ones. However, when individual types of problem are considered, students seemed to have more difficulties in learning stress and intonation than in other aspects. It is notable that the majority of participants who reported having this type of problem belonged to group BUH1, who had not had any formal instruction by the time of the study. This group also showed concerns in more areas of pronunciation than the other two groups, which may be due to the same reason.

Table 2

Learning difficulties

Problem		BUH1	BUH2	VUS	Total
Supra-segmental	Stress & Intonation	7	3	1	11
	Connected speech	2	0	0	2
Segmental	Final sounds	3	2	2	7
	Confusing sounds	1	3	2	6
	Different inventories of sounds	2	0	1	3
	Vowels	1	1	0	2
	Silenced sounds	0	1	0	1
Instruction	Lack of knowledge of phonetic transcription	2	0	1	3
	Lack of corrective feedback	1	0	1	2
Others	Existence of many native accents	2	0	0	2
	Pronunciation of sentences	1	1	0	2
	Effects of regional Vietnamese accents	2	0	0	2
	Control of articulators	1	0	0	1

Not many participants were able to explain why they encountered the problems; however, a number of issues did come up. They are summarized in Table 3.

Table 3

Possible causes of learning difficulties

Difficulty		Possible cause
Supra-segmental	Stress & Intonation	Too many patterns Differences between English and Vietnamese No understanding of stress placement Lack of feedback Bad habits Failure to use in real life despite awareness of rules
	Connected speech	Unfamiliarity
Segmental	Final sounds	Non-existence in Vietnamese Unnatural (exaggerated) production of final sounds Awareness of rules but lack of automaticity or habits
	Confusing sounds	Inability to perceive and produce them differently correctly
	Different inventories of sounds	Problems with sounds not existing in Vietnamese
	Vowels	Differences between English and Vietnamese vowels
	Silenced sounds (e.g. ***h***our, recei***p***t)	Failure to remember all cases
Instruction	Lack of knowledge of phonetic transcription	No instruction available in school
	Lack of corrective feedback	(No further comments)
Others	Existence of many native accents	Confusion
	Pronunciation of sentences	Distraction by effort to produce long utterances Difficulty in controlling articulators
	Effects of regional Vietnamese accents	Heavy impact on production of English speech
	Control of articulators	(No further comments)

Evaluations and expectations of the current teaching

When asked whether taking a pronunciation course at school was beneficial, 37 participants responded positively and identified various benefits of receiving instruction. The three groups shared some opinions of the values of pronunciation training, especially its effect on improving understanding and communication. More importantly, group BUH2 insisted that they had made a

lot of improvements after taking the course. Table 4 lists all the benefits that the participants mentioned and indicates by which group each benefit was perceived.

Table 4

Benefits of taking a pronunciation course

Benefit	BUH1	BUH2	VUS
Lack of practice during high school; the need for more practice	X		X
More focus on pronunciation		X	X
Improvements in speaking skills			X
Improvements in pronunciation skills	X		
Improvements in comprehension and communication	X	X	X
Difficulty in self-study	X		
Biggest weakness of Vietnamese learners	X		

Finally, the participants were asked for their views of the teaching practices they were experiencing or were about to do so. For group BUH1, who had not taken the compulsory pronunciation training course, the question was: *What do you expect the teachers to teach and how do you expect them to teach it?* They seemed to have high expectations of the teachers when looking for their standard pronunciation, use of a variety of activities and exercises and provision of corrective feedback. They would like to learn about the IPA, the contrast between Vietnamese and English sounds and different accents of English. They also expected their teachers to motivate them to learn and help them get rid of their Vietnamese regional accents.

For group BUH2, who had already taken the course, the question was: "*What recommendations would you make to the course you have taken?*" Their responses, to some extent, reveal the problems they encountered during the course. First, despite an intensive focus on pronunciation, instruction delivered over a relatively short period of time (nine weeks) and only once a week was not enough for them to make significant improvements. In fact, they reported that there were too many things to learn in such a short course. Second, they suggested that more practice be given outside class. Finally, they expected to be inspired more by the teachers and to have contact with native speakers.

The conversations with the participants in group VUS brought even more diverse responses when I encouraged them to talk about what was happening in their classes as regards pronunciation teaching. Some common issues arose from their comments. First, they shared with the other two groups the same concern about their non-native teachers' accents. Second, pronunciation seemed to receive little attention from the teachers, several of whom were reported to skip the pronunciation activities in the textbook. As a result, some learners claimed that they did not have any activities either in class or at home for practice and many others were even confused between learning speaking and learning pronunciation.

DISCUSSION

Learning goals

Findings show that Vietnamese learners of English still aim at achieving native-like accents. In this case, the nativeness principle seems to dominate the current English teaching and learning in Vietnam. More importantly, in a context where learners still consider native-like pronunciation a proof of language competence and success in learning, this goal may become a burden for both the teachers and learners. In addition, identity appears to have no effect on learning as only two students expressed the fear of losing their identity.

It is interesting to find that, in aiming for nativeness, these learners preferred an American accent to a British accent as a model for learning. However, a closer look at the reasons for such a selection will confirm Setter (2008)'s assumption that it is a result of sociocultural, political and market-driven choices. Some participants did explain that at the present time, an American accent is needed to get a job, that many textbooks available are in American English, and that there are a lot of American movies on TV and the Internet.

Learning difficulties

Findings reveal that learners' problems range from issues related to the subject itself and learners' background to external factors such as instruction. The long-existing debate over what aspect to teach, segmentals or suprasegmentals, seems to remain inconclusive in this context, as the participants reported having difficulties in both areas. Moreover, instruction seemed to have certain effects on their learning as groups BUH2 and VUS altogether raised fewer issues than group BUH1. Especially, group BUH2, who had already received formal training in pronunciation, did not face the instruction-related issues as did the other two groups.

Further discussion on the causes of learning difficulties has brought in several interesting insights. First, L1 transfer plays some role in the acquisition of L2 and influences learners' performance (Navehebrahim, 2012; Thomas, 2013). In this case, it is a source of learning difficulties. Second, these learners appeared to be lacking in necessary skills such as listening (Barreiro, 2002; Couper, 2011; Rogerson-Revell, 2011), discrimination (Gilakjani & Ahmadi, 2011; Kissling, 2014), and even motor skills (Hockett, 1950). Finally, the fact that learners could not remember to pronounce final sounds or use intonation appropriately in spontaneous speech contexts despite an awareness of the rules may indicate that there was conscious learning (Krashen, 1982) but not automaticity.

Evaluations and expectations of the current teaching

It is remarkable that the contexts in which the three groups perceived the value of instruction were different due to the availability of instruction as well as the manner in which it was conducted. Group BUH1 had not had any training, so the benefits that were named were mostly assumed while those recognized by group BUH2 were ones they actually received after completing the course. In another context, learners from VUS had received little instruction due to the approach their teachers used to integrate pronunciation into their classrooms. Interestingly, despite the different contexts, these learners highly valued and thus requested for more instruction.

Shared beliefs can also be seen from the results of the last questions. First, teachers were expected to have very good pronunciation and native speakers were still considered good models. Second, there seemed to be not enough time for them to learn and practice in class. Finally, teachers needed to motivate learners through a variety of tasks and activities as well as provide corrective feedback.

The biggest difference is in the area of teaching contents, which, once again, may be due to the amount of and the way in which instruction was given.

CONCLUSION

The present study has shown that Vietnamese adult learners of English desire to obtain native-like pronunciation. In learning the new phonological system, they often encounter problems with stress and intonation as well as final sounds. Instruction is perceived to be highly useful, but current teaching practices have not met their expectations in terms of the teacher's accent, length and focus of instruction, types of activities used and provision of feedback.

In addition, this research also has a number of implications for future improvements in the current teaching and learning of English pronunciation in Vietnam. First, it is vital that teachers themselves be cognizant of the intelligibility principle, consider it the new target for teaching and then adjust their teaching approach to suit it. Murphy (2014) claims that it is unfair and unethical for teacher to make their learners believe that they will ever be able to achieve native-like pronunciation. So, teachers also need to make learners aware of a more realistic and achievable goal for learning: being intelligible. This awareness will help lessen the burden and pressure placed on both teachers and learners and foster learners' confidence so that they can get better outcomes.

Second, learners have a variety of difficulties in learning pronunciation, one of which is the failure to perform effectively in spontaneous or real life situations despite their knowledge of the forms and success performance in classroom activities. What else is needed to bridge the gap between conscious learning and automaticity? This question remains an ever-important mission for L2 researchers and teachers.

Third, from learners' perspectives, pronunciation instruction is important not only because it helps to improve their pronunciation. "Students may … see a pronunciation course as valuable for various reasons, including increased awareness, improved listening skills, motivation, and perceived improvement" (Henrichsen & Stephens, 2015, p. 202). L2 teachers, therefore, need to be aware of the many benefits of instruction in order to gain more confidence and stop neglecting it in their classrooms.

Fourth, learners' needs are important as they will define their learning targets, shape their motivation and determine whether or not learning will sustain outside class. Teachers, therefore, ought to be aware of their needs and provide support accordingly for successful learning to happen.

Finally, findings show that learners, without any knowledge of language learning or teaching methodology, know quite clearly what problems they are facing and what they need for improvement. Therefore, it is worth involving students in planning, delivering and evaluating pronunciation lessons.

ABOUT THE AUTHOR

Duc Nguyen Anh Dao is currently a PhD student at the School of Education, University of Nottingham, Malaysia Campus. Before starting her study in Malaysia, Duc was a lecturer at Banking University of Ho Chi Minh City, Vietnam. She holds a Master in TESOL Studies from the University of Queensland, Australia and an MBA from Bolton University, UK. Her research interests are pronunciation, phonetics, phonology and teaching methodology.

Email: kabx6dna@nottingham.edu.my

Mobile: +84-909381626

REFERENCES

Alghazo, S. M. (2015). Advanced EFL learners' beliefs about pronunciation teaching. *International Education Studies*, *8*(11), 63. https://doi.org/10.5539/ies.v8n11p63

Barreiro, S. C. (2002). Speech perception in L2. *Odisea*, (2), 7–14.

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.

Couper, G. (2011). What makes pronunciation teaching work? Testing for the effect of two variables: socially constructed metalanguage and critical listening. *Language Awareness*, *20*(January 2015), 159–182. https://doi.org/10.1080/09658416.2011.570347

Cunningham, U. (2009). Phonetic correlates of unintelligibility in Vietnameseaccented English. In *FONETIK 2009 The XXIIth Swedish Phonetics Conference*.

Cunningham, U. (2013). Teachability and learnability of English pronunciation features for Vietnamese-speaking learners. In E. Waniek-Klimczak & L. R. Shockey (Eds.), *Teaching and Researching English Accents in Native and Non-native Speakers* (pp. 3–14). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-24019-5

Darcy, I., Ewert, D., & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL teachers' pronunciation "toolbox." In J. Levis & Levelle K. (Eds.), *Proceedings of the 5th pronunciation in second language learning and teaching conference* (pp. 93–108). Ames: IA: Iowa State University.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *Source: TESOL Quarterly*, *3938194*(3), 379–397. Retrieved from http://www.jstor.org/stable/3588486

Gilakjani, A. P., & Ahmadi, M. (2011). Why is pronunciation so difficult to learn? *English Language Teaching*, *4*(3), p74. https://doi.org/10.5539/elt.v4n3p74

Ha, C. T. (2005). Common pronunciation problems of Vietnamese learners of English. *Journal of Science - Foreign Languages*. Retrieved from http://js.vnu.edu.vn/Ngoaingu_1/Bai3.pdf

Henrichsen, L., & Stephens, C. (2015). Advanced adult ESL students' perspectives on the benefits of pronunciation instruction. In J. Levis, R. Mohammed, M. Qian, & Z. Zhou (Eds.), *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching Conference, Santa Barbara, CA* (pp. 197–205). Ames: IA: Iowa State University.

Hoang, V. Van. (2010). The current situation and issues of the teaching of English in Vietnam. *International Symposium on the Teaching of English in Asian (2)*, *22*(1), 7–18. Retrieved from http://www.ritsumei.ac.jp/acd/re/k-rsc/lcs/kiyou/pdf_22-1/RitsIILCS_22.1pp.7-18_HOANG.pdf

Hockett, C. F. (1950). Learning pronunciation. *The Modern Language Journal*, *34*(4), 261–269.

Ketabi, S., & Saeb, F. (2015). Pronunciation teaching : Past and present. *International Journal of*

Applied Linguistics & English Literature, *4*(5), 182–189. https://doi.org/10.7575/aiac.ijalel.v.4n.5p.182

Kissling, E. M. (2014). What predicts the effectiveness of foreign-language pronunciation instruction? Investigating the role of perception and other individual differences. *Canadian Modern Language Review/ La Revue Canadienne Des Langues Vivantes*, *70*(4), 532–558. https://doi.org/10.3138/cmlr.2161

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Hayward, CA: *Alemany Press*

Lear, E. (2011). Using guided reflective journals in large classes: Motivating students to independently improve pronunciation. *The Asian EFL Journal Quarterly*, *13*(September), 113–137.

Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation. *TESOL Quarterly*, *39*(3), 369–378. https://doi.org/10.2307/3588485

Moghaddam, M. S., Nasiri, M., Zarea, A., & Sepehrinia, S. (2012). Teaching pronunciation: The lost ring of the chain. *Journal of Language Teaching and Research*, *3*(1), 215–219. https://doi.org/10.4304/jltr.3.1.215-219

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, *25*(3), 481-520. https://doi.org/10.2307/3586981

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *45*(1), 73–97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System*, *42*, 258–269. https://doi.org/10.1016/j.system.2013.12.007

National Foreign Languages 2020 Project. (n.d.). Retrieved from http://dean2020.edu.vn/vi/

Navehebrahim, M. (2012). An investigation on pronunciation of language learners of English in Persian background: Deviation forms from the target language norms. *International Conference on Education & Educational Psychology (ICEEPSY 2012)*, *69*(Iceepsy), 518–525. https://doi.org/10.1016/j.sbspro.2012.11.441

Nguyen, N. (2002). Interlanguage phonology and the pronunciation of English final consonant clusters by native speakers of Vietnamese, 1–20. Retrieved from http://www.ohio.edu/linguistics/workingpapers/2008/nguyen_2008.pdf

Nguyen, T. H. (1998). Difficulties of Vietnamese learners with the length of English sounds when they produce and suggested strategies. *Tạp Chí Khoa Học & Công Nghệ*, *112*(12), 231–235.

Nguyen, T. H. (2002). Some proposed measures to help Vietnamese learners confront with the English final consonant clusters. *Tạp Chí Khoa Học & Công Nghệ*, *91*(3), 71–75.

Osburne, A. G. (1996). Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics*, *17*(2), 164–181. https://doi.org/10.1093/applin/17.2.164

Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. London : Continuum.

Setter, J. (2008). Theories and approaches in English pronunciation. In R. Monroy & A. Sanchez (Eds.), *25 Years of Applied Linguistics in Spain: Milestones and challenges. Murcia: Universidad de Murcia de Publications* (pp. 447–457). Murcia: Universidad de Murcia, Servicio de Publacaciaones.

Thomas, M. (2013). History of the study of second language acquisition. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge Handbook of Second Language Acquisition*. Cambridge: Cambridge University Press.

Vo, H., & Hoai, T. (2016, November 16). Bộ trưởng Giáo dục: Đề án ngoại ngữ hơn 9.000 tỷ không đạt mục tiêu. *VnExpress*. Retrieved from http://vnexpress.net/tin-tuc/giao-duc/bo-truong-giao-duc-de-an-ngoai-ngu-hon-9-000-ty-khong-dat-muc-tieu-3499623.html

APPENDIX – INTERVIEW QUESTIONS

1. Would you like to be mistaken for a native speaker? Why?
2. What do you think is the best English speaker pronunciation model? Why?
3. Do you have any difficulties in learning English pronunciation? Is yes, what are they?
4. What may be the cause of those problems?
5. Do you think taking a pronunciation course is beneficial?
6. a. What do you expect the teachers to teach and how do you expect them to teach it?
   b. What recommendations would you make to the course you have taken?
   c. What do you think about the way pronunciation is being taught in class?

Huensch, A. (2018). Examining foreign language instructors' definitions of pronunciation instruction. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 100-110). Ames, IA: Iowa State University.

EXAMINING FOREIGN LANGUAGE INSTRUCTORS' DEFINITIONS OF PRONUNCIATION INSTRUCTION

Amanda Huensch, University of South Florida

As part of a larger investigation exploring the training, beliefs, and practices of foreign language teachers (i.e., teachers of languages other than English), the current study examined how instructors defined pronunciation instruction and their reports of how their textbooks approached pronunciation instruction. Participants included 296 beginning-level foreign language instructors of French, German, and Spanish from large (over 15,000), public universities in the US who completed an online survey via Qualtrics. Results indicated that the most common themes from instructors' definitions of pronunciation instruction (i.e., focusing on sounds and using listen and repeat activities) were the same as those reported in how their textbooks approached pronunciation. In addition, results indicated that fewer instructors' definitions of pronunciation instruction included a mention of potential learning outcomes, and those that did comment on learning outcomes more often mentioned a goal of native-like pronunciation compared to goals of intelligible/comprehensible speech. Findings highlight the potentially influential role materials have in shaping instructors' conceptualization of pronunciation instruction and provide direct implications for materials developers to integrate pronunciation foci targeting intelligible communication as a goal.

INTRODUCTION

An important area of pronunciation research is that which explores the classroom practices and teaching beliefs of language instructors. This research is important because it illuminates what is occurring in language classrooms and can both provide (1) an understanding of how well research findings are making their way into language classrooms and (2) guidance for the creation and modification teacher training opportunities and teaching materials. While a relatively substantial body of literature exists investigating instructors in ESL/EFL contexts (e.g., Breitkreutz, Derwing, & Rossiter, 2001; Buss, 2016; Foote, Holtby, & Derwing, 2011; Foote, Trofimovich, Collins, & Urzúa, 2016; Kirkova-Naskova et al., 2013; Tergujeff, 2012), fewer studies (e.g., Delicado Cantero & Steed, 2015; Huensch, 2018) have investigated foreign language (FL) contexts, or those in which a language other than English is taught. But with millions of students enrolled in FL classes in the US and research suggesting that these learners think it is important to improve their pronunciation (Huensch & Thompson, 2017), instructors' beliefs and practices in these contexts should be better understood.

Several previous investigations of instructors' beliefs and classroom practices related to pronunciation have collected data using a survey first introduced in Breitkreutz et al. (2001). These include Breitkreutz, Derwing, & Rossiter, 2001 and Foote et al. (2011) in the Canadian ESL context, Buss (2016) in the Brazilian EFL context, and Huensch (2018) in the US FL context. These investigations focused on the amount and type of pronunciation instruction being delivered in language classes; instructors' prior training in pronunciation teaching (and language teaching in general); and instructors' beliefs related to the goals and importance of teaching pronunciation. Results for ESL/EFL and FL contexts have indicated that pronunciation does not

represent a major focus of language courses, which has been corroborated by classroom observation research in an ESL context (Foote et al., 2016). When asked to identify the most serious pronunciation problems faced by students, ESL instructors indicated a mix of both segmental and suprasegmental features whereas EFL and FL instructors tended to provide more mentions of segmental features. Regarding training, results from Foote et al. (2011) and Buss (2016) for ESL/EFL contexts additionally indicated that only about one third of instructors reported having taken a course focused specifically on L2 pronunciation teaching, but that most (e.g., over 90%) had formal TEFL/TESL training. In comparison, in the FL context, while similar to ESL/EFL contexts a minority of instructors indicated having taken a course focused specifically on L2 pronunciation teaching (13%), a striking difference was found regarding formal FL training such that 50% reported no formal FL teaching certification. Huensch attempted to explain these findings by pointing out that instructors of introductory-level FL courses typically have less teaching experience (two thirds of instructors in her study were graduate teaching assistants).

The fact that instructors in FL contexts are more likely to be novice teachers heavily reliant on teaching materials provided to them such as their course textbooks, understanding the place of pronunciation in those textbooks becomes a question of central concern. Previous research exploring how pronunciation is incorporated in introductory FL textbooks has been conducted with both Spanish (Arteaga, 2000) and German (Pittman, 2015) textbooks. Both studies indicated that pronunciation was not well-represented. For example, Pittman reported that five of the top ten most used German textbooks included no mention of pronunciation. Arteaga reported similar results for 10 introductory Spanish textbooks, and additionally reported that even when pronunciation was included, its presentation often lacked sufficient recycling, included inaccurate or confusing explanations, and did not assist students in developing self-monitoring capabilities. Missing from these textbook investigations, however, is information from instructors about their implementation and reception, which might be potentially enlightening given the lack of training of many instructors in these contexts. Therefore, the current study reports on FL instructors' textbook use and their understanding of how pronunciation instruction is presented.

A final focus of previous research investigating instructors' beliefs and classroom practices has been related to instructors' reported beliefs about the importance of teaching pronunciation and what the goals for learning should be. All of the studies that have used the Breitkreutz survey, regardless of whether in an ESL/EFL or FL context, indicated that a majority of instructors think it is important to teach pronunciation (Breitkreutz et al., 2001; Buss, 2016; Foote et al., 2011; Huensch, 2018). Regarding goals for instruction, of potential interest is whether goals are related to nativeness or intelligibility principles (Levis, 2005)—the former emphasizing native-like pronunciation and the latter focusing on being understood despite the presence of a foreign accent. In recent years, researchers have argued for a prioritization of intelligibility and comprehensibility (e.g., Derwing & Munro, 2015, p. 9-10) and a question has been whether this prioritization is also present in language classrooms. In terms of the FL context specifically, Huensch demonstrated that while instructors strongly agreed that the goal of pronunciation teaching should be to increase learners' intelligibility, they did not as strongly disagree that accent reduction should not be a goal. In addition, when discussing assessment practices instructors indicated that pronunciation was not included because the course focused on communicative skills. The author argued that findings indicated that some instructors conceptualized of pronunciation instruction from a nativeness paradigm, perhaps viewing

pronunciation instruction as synonymous with accent reduction. Because of this definition, they thus regarded pronunciation teaching as incompatible with communicative course goals. In order to shed light on this question, the current study directly investigated how FL instructors define and conceptualize pronunciation instruction.

Based on the findings from previous research, the current study investigated two specific areas related to the beliefs and practices regarding pronunciation of beginning-level instructors of French, German, and Spanish from US universities. Instructors from this context were chosen both because explorations of foreign languages (as opposed to ESL/EFL) are underrepresented in the field (Thomson & Derwing, 2015) and because this population teaches millions of language learners in the US. The first goal of the study was to better understand how instructors define and conceptualize of pronunciation instruction. The second goal was to explore instructors' perceptions of how pronunciation instruction is approached in the textbooks they use. Following are the research questions.

Research Questions

1. What are the most frequent themes that emerge in beginning-level French, German, and Spanish instructors' definitions of the term 'pronunciation instruction'?
2. What are the most frequent themes that emerge when beginning-level French, German, and Spanish instructors describe the approach to pronunciation instruction taken by their textbooks?

METHODS

Participants

Instructors teaching basic language courses (the first four semesters) in French, German, and Spanish from 27 large (over 15,000 students), public universities in the US were identified via web searches and recruited via email to participate. Data in the current study come from the same set of participants as in Huensch (2018) and included 296 instructors: French (n = 89), German (n = 80), and Spanish (n = 127). While participants had a range of teaching experience and qualifications, a majority were graduate teaching assistants (n = 197) who were teaching first or second semester classes (n = 187). Participants received a $10 Amazon gift card for completing the survey.

Materials and Procedure

The online survey was adapted from the instrument used in Foote et al. (2011) and included 81-103 questions (depending upon whether additional questions were triggered as follow-ups), took approximately 30 minutes to complete, and was delivered via Qualtrics. Response types included multiple-choice, open-ended responses, and Likert-scale ratings. The complete survey is available on IRIS.

Data Analysis

Data in the current study come from multiple-choice (e.g., "Do you use pronunciation activities regularly or sporadically?") and open-ended response (e.g., "What should your textbook do differently, if anything, to improve its approach to teaching pronunciation?") questions. Responses from the multiple-choice questions are reported as percentages, separated by language

group (French, German, and Spanish) to allow for comparison but also reported as a total percentage for the entire 296 instructors. Percentages are used because the sample size in each language group was not the same. In analyzing the open-ended responses, the author and a research assistant each coded all of the data based on a rubric developed by the author which was compiled based on emergent themes. These codings were compared and any instances of disagreement were discussed and resolved. The coded responses were then tallied for frequency to determine the most common themes among the instructors' responses; themes that were reported by greater than 10% of the instructors were included in the current analysis. Representative quotations from the open-ended responses were then selected and are reported to illustrate the themes.

RESULTS

Research question one explored how instructors of beginning-level language courses defined pronunciation instruction. The first prompt given to participants when completing the online survey was the following: "Please take a moment to think about what the phrase 'pronunciation instruction' means to you in the context of foreign language teaching. In the space below, please briefly explain how you define pronunciation instruction and provide a few examples that come to mind." Emergent themes from the responses were related to three broad themes: (a) course content (i.e., what is being taught such as segmentals/sounds, sound-spelling correspondences), (b) instructional techniques (i.e., how the course content is being taught such as using corrective feedback, listen and repeat), and/or (c) learning outcomes (i.e., what outcomes would be desirable such as increased intelligibility, improved listening skills). Across all three categories and languages, the most frequently mentioned element was a course content theme: 64% of the definitions (64% French, 64% German, 65% Spanish) made reference to the teaching/learning of sounds/segments/consonants and vowels. The second most frequently mentioned element was an instructional technique: 33% of the definitions (43% French, 40% German, 21% Spanish) mentioned using a listen and repeat technique. Figures 1, 2, and 3 indicate the themes that 10% or more of the instructors mentions separated by language group.

Figure 1 summarizes the results for the themes related to course content and demonstrates that after sounds and words, the most frequently mentioned items were related to making cross-linguistic comparisons and teaching about sound-spelling correspondences.

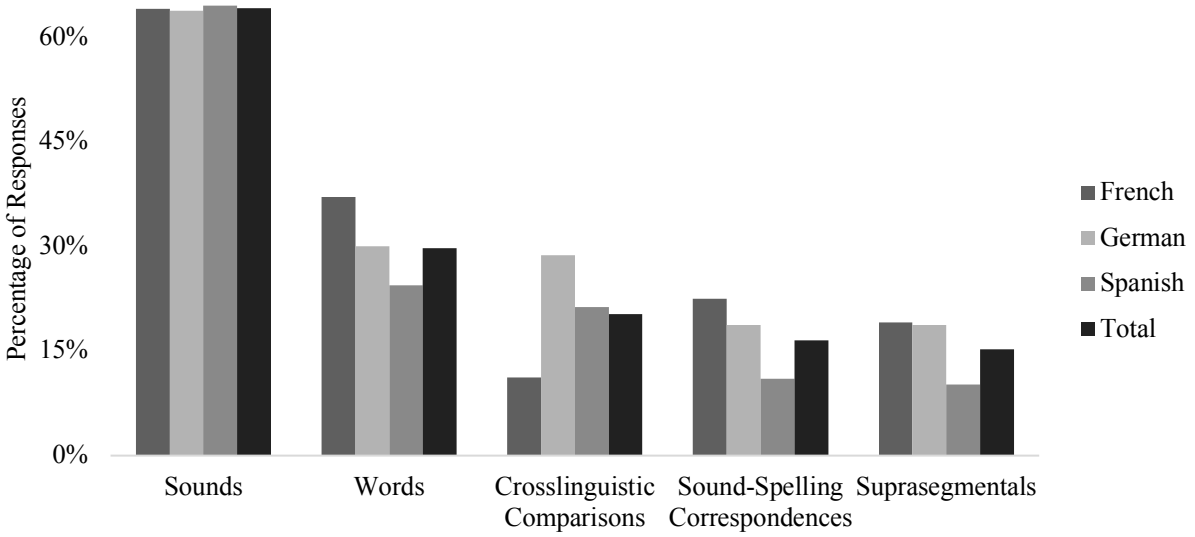


Figure 1. Course content mentioned in instructors' definitions of pronunciation instruction.

In definitions (1) and (2) from French instructors, for example, mentions of the 'phonetic system' and of 'particular sounds' were coded as *sounds*. An example of a statement coded as cross-linguistic comparison can be seen in definition (2) where the instructor mentions 'making parallels with similar sounds in the student's native language'. Results were relatively similar across the language groups, except perhaps that the French instructors were less likely to mention cross-linguistic comparisons. For all the groups, suprasegmentals were the least often mentioned.

(1) French instructor

> *Pronunciation instruction is helping students to understand* ***the phonetic system of a language*** [sound] *or the* ***pronunciation of a specific word*** [word]*. Sometimes I have my students* ***repeat words after me*** [listen and repeat] *as a group or individually. Other times we have discussed the unique vowels that French has and* ***how to both pronounce and recognize them*** [listening discrimination].

(2) French instructor

> *Giving clear example of how to pronounce* ***particular sounds*** [sound]*, making parallels with similar sounds* ***in the student's native language*** [cross-linguistic comparison]*,* ***providing constant feedback*** [corrective feedback]

Figure 2 summarizes the results for the instructional techniques focus and demonstrates that the most common mentions after 'listen and repeat' were using corrective feedback, providing an articulatory description of a sound, and acting as a model in the classroom.

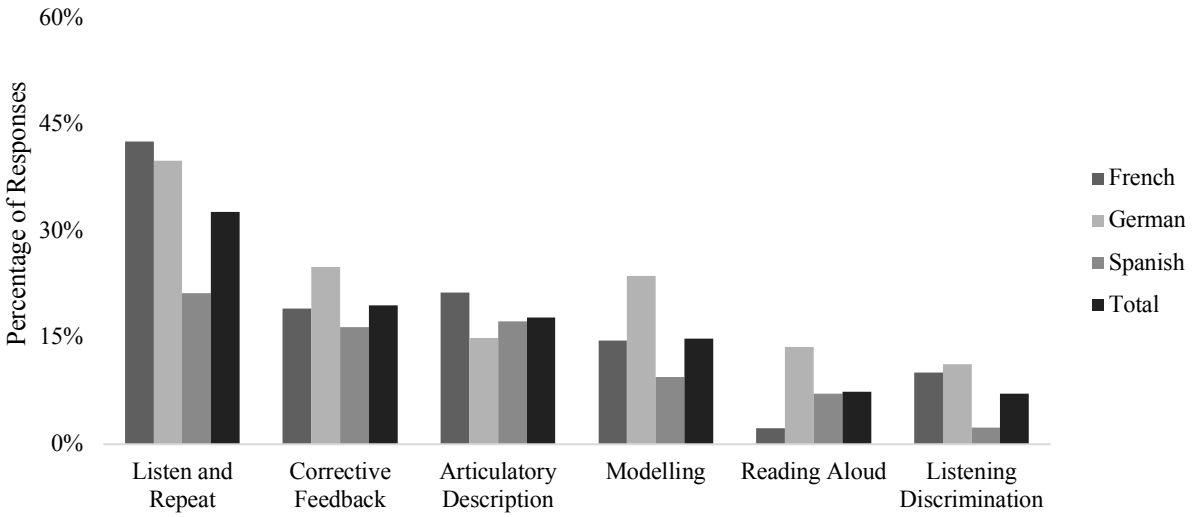


Figure 2. Instructional practices mentioned in instructors' definitions of pronunciation instruction.

Definitions (2) and (3) provide examples that were coded as corrective feedback with their mentions of 'providing constant feedback' and 'correcting students' pronunciation'. Definition (3) includes mentions of modelling, listen and repeat, and articulatory description.

(3) German instructor

> *Pronunciation instruction means teaching students to pronounce a foreign language so that their pronunciation at least* ***does not impede comprehension*** [intelligible] *and* ***ideally is that of a native speaker*** [native-like pronunciation]. *Pronunciation instruction includes, for example,* ***modelling native pronunciation*** [modelling]*,* ***correcting students' pronunciation*** [corrective feedback] *(e.g. during* ***reading*** [reading aloud] *or communication exercises), and also specific targeted exercises, such as tongue twisters and word pairs/groups to practice pronunciation.*

Figure 3 summarizes the results for the learning outcomes focus and demonstrates that learning outcomes were mentioned overall less than course content and instructional techniques. Approximately 20% of responses mentioned of 'correct'/native-like pronunciation, as shown in definitions (3) and (4) whereas approximately 10% of responses mentioned intelligible/comprehensible pronunciation, as shown in definitions (3) and (5).

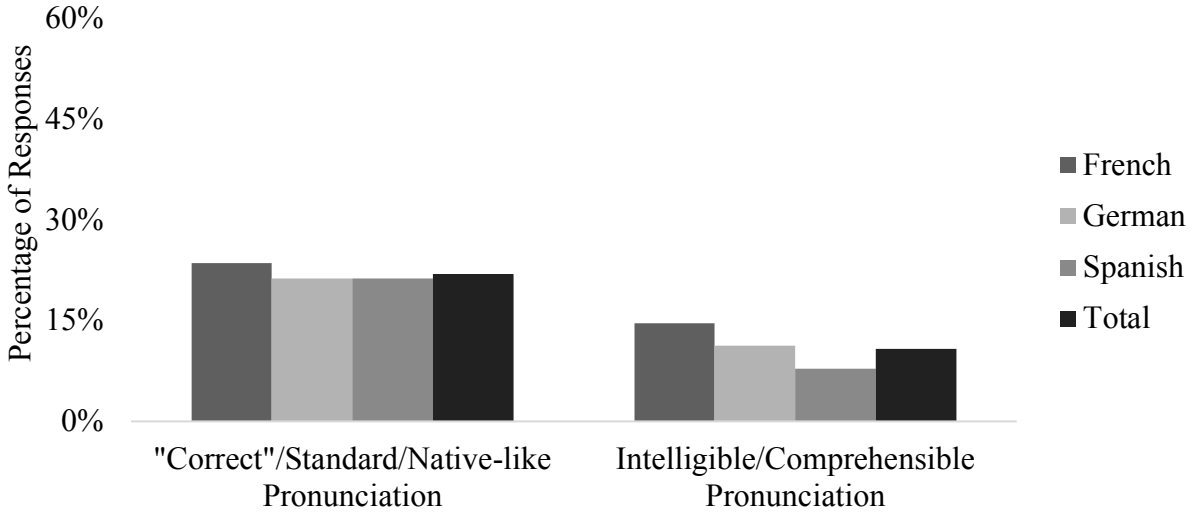


Figure 3. Learning outcomes mentioned in instructors' definitions of pronunciation instruction.

(4) Spanish instructor

Training a secondary language learner to speak and ***enunciate words*** [words] ***as closely as possible to a native speaker*** [native-like pronunciation]. *-****Explicit correction*** *via correct pronunciation* [corrective feedback]. *-Rephrasing into correct pronunciation. -Exposure/induction by* ***listening to authentic audio*** *in the target language* [modelling].

(5) Spanish instructor

Pronunciation Instruction makes me think of how we teach the production of the ***sounds*** [sound] *of a language with input and other goal-oriented methods so the speaker is* ***comprehensible*** [comprehensible] *and fluent in a language.*

Research question two explored how instructors characterized their textbooks' approaches to pronunciation instruction. Instructors were first asked whether they used any of the pronunciation activities from their textbook. Twenty-eight percent (22% French, 28% German, 32% Spanish) reported that their textbooks did not include pronunciation activities. Forty-three percent (54% French, 49% German, 32% Spanish) reported using pronunciation activities. The 128 instructors who reported using the activities were then asked whether they use the activities regularly or sporadically, and an open-ended question about the approach to pronunciation teaching taken by the textbook. Forty-two percent (56% French, 36% German, 32% Spanish) reported using the pronunciation activities regularly. Regarding instructors' characterization of their textbook's approach to pronunciation instruction, the 125 instructor comments showed similarities to the definitions of pronunciation (see the Appendix for the complete data table). For example, similar to the definitions, mentioning a focus on sounds was the most frequent theme (44%; 51% French, 44% German, 36% Spanish) and the technique of listen and repeat was the second most common (26%; 27% French, 31% German, 21% Spanish). The next most common mentions were related to the content or delivery of the textbook with 20% (16% French, 21% German, 24% Spanish) mentioning the inclusion of audio/video recordings and 18% (13% French, 5% German, 33% Spanish) identifying the homework/practice as being online. However, in the case of the latter, the result appears to be driven by the Spanish instructor responses. A

theme that emerged that was not present in the pronunciation instruction definitions was that 10% (11% French, 8% German, 12% Spanish) indicated that the textbook connected pronunciation with vocabulary learning, as in comment (6).

(6) Spanish instructor

*Charts with explanations of the **sounds*** [sound] *and letters of some words specially difficult for **English native speakers*** [cross-linguistic comparison]. *Also it encourages us to pronounce carefully and focused on pronunciation the **vocabulary section*** [vocab] *through **repetitions and make students repeat*** [listen and repeat]

Another theme included highlighting the short or incidental nature of the textbook's approach as in comment (7).

(7) French instructor

It tries to teach it as a "by the way" phenomenon. "Just for your information"

There were very few mentions related to learning outcomes and none of them reached 10%. Comments (8) and (9), however, both reference communicative language teaching, but comment (8) explains that pronunciation is not incorporated because the class uses a communicative method while comment (9) explains how pronunciation is incorporated communicatively.

(8) Spanish instructor

it is a communicative method, so I think the emphasis is not on the pronunciation, however, I use repetition when my students cannot say a word accurately.

(9) French instructor

CLT; pronunciation is incorporated into activities surrounding a film that the students watch throughout the semester. They are asked to describe their reactions to the film, and each section focuses on particular phonetic pronunciation in French. Their responses and vocabulary surrounding their responses emphasize these pronunciation lessons.

DISCUSSION AND CONCLUSION

The first research question explored FL instructors' conceptualization of pronunciation instruction via definitions and examples they provided of the concept. Results indicated that a majority of responses included mentions of focusing on sounds as course content and using an instructional technique of listen and repeat. The finding that sounds were more often mentioned that suprasegmental features corroborates previous work in EFL/FL contexts (Buss, 2016; Huensch, 2018) in which instructors more often indicated that difficult problems for students were segmental rather than suprasegmental. Of interest in the current study was how, if at all, instructors discussed course goals from either nativeness or intelligibility principles. Fewer instructors overall indicated learning outcomes in their definitions as compared to course content or instructional practices. When they did, they more often referred to a focus on 'correct' or native-like pronunciation (22%) than intelligible/comprehensible pronunciation (11%). While this result speaks to how instructors conceptualize of pronunciation instruction, it does not necessarily represent their personal teaching beliefs. In other words, it is possible that they believe course goals should be communicative and that pronunciation instruction as a practice entails a goal of native-like pronunciation. This perhaps explains statements such as (8) in which

the instructor explains a lack of pronunciation focus by indicating that their course adopts a communicative method. Thus, a first step in providing better pronunciation instruction training for instructors should include a discussion of the role pronunciation plays in intelligible communication and practical ideas for integrating pronunciation instruction into activities with a communicative focus.

The second research question investigated how textbook approaches to pronunciation are characterized by instructors. Instructors' comments about their textbooks echoed the findings of previous research that directly examined pronunciation foci in textbooks (Arteaga, 2000; Pittman, 2015). Results also indicated that there was a parallel in the findings from research questions one and two in that the most commonly mentioned themes were a focus on sounds and the use of a listen and repeat instructional technique. One possible explanation for this finding could be connected to the fact that instructors of beginning-level FL courses are more often novice teachers with little training or background language teaching, generally, and pronunciation teaching, specifically. Thus, their conceptualization of pronunciation instruction may well be influenced by how teaching materials approach the topic. These findings also speak to the need for better pronunciation instruction training for foreign language teachers, which includes developing skills for evaluating instructional materials and practice in supplementing those materials.

While this study has provided some insights into pronunciation instruction in beginning-level FL classes, several limitations should be acknowledged. First, as data were gathered via an online questionnaire, the findings might not necessarily be representative of teachers' actual practices and they might not accurately reflect the textbook content. The use of additional data collection methods such as classroom observations and updated investigations of textbooks would be beneficial. In addition, the study was limited to beginning-level classes (i.e., those at the first four semesters of instruction). Given that it is typical for FL degrees to offer a phonetics class as an upper-level required course or elective, understanding the goals and materials used in these courses would also be beneficial. Despite these limitations, this study has provided a better understanding of an understudied population: foreign language instructors of beginning-level language courses. This study also provides directions for future work. An important next step is designing and testing teacher training materials to help instructors integrate pronunciation foci targeting intelligible communication as a goal.

ACKNOWLEDGMENTS

This project was funded by a University of South Florida New Research Grant. I am grateful to Jennifer Foote, Amy Holtby, and Tracey Derwing, who allowed me access to their survey to adapt it for the current project. I would also like to thank the participants and my research assistant, Aneesa Ali.

ABOUT THE AUTHOR

Amanda Huensch (Ph.D. University of Illinois at Urbana-Champaign, 2013) is Assistant Professor of Applied Linguistics in the Department of World Languages at the University of South Florida, where she teaches courses in second language acquisition, pronunciation pedagogy, and applied linguistics. Her research examines second language speech development

in and outside of the classroom. Her most recent work has been published in *The Modern Language Journal*, the *Journal of Second Language Pronunciation*, and *Applied Psycholinguistics*.

REFERENCES

Arteaga, D. L. (2000). Articulatory phonetics in the first-year Spanish classroom. *The Modern Language Journal*, *84*, 339–354.

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

Buss, L. (2016). Beliefs and practices of Brazilian EFL teachers regarding pronunciation. *Language Teaching Research, 20,* 619–637.

Delicado Cantero, M., & Steed, W. (2015). La enseñanza de la pronunciación del español en Australia: creencias y actitudes de los profesores. *Journal of Spanish Language Teaching, 2,* 18–35.

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal*, *29*, 1–22.

Foote, J. A., Trofimovich, P., Collins, L., & Urzúa, F. S. (2016). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal*, *44*, 181–196.

Huensch, A. (2018). Pronunciation in foreign language classrooms: Instructors' training, classroom practices, and beliefs. *Language Teaching Research*, 1–20. DOI: 10.1177/1362168818767182

Huensch, A., & Thompson, A. S. (2017). Contextualizing attitudes toward pronunciation: Foreign language learners in the US. *Foreign Language Annals, 50*, 410–432.

Kirkova-Naskova, A., Tergujeff, E., Frost, D., Henderson, A., Kautzsch, A., Levey, D., Murphy, D., & Waniek- Klimczak, E. (2013). Teachers' views on their professional training and assessment practices: Selected results from the English Pronunciation Teaching in Europe survey. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 29–42). Ames, IA: Iowa State University.

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

Pittman, I. (2015). Integration of pronunciation in first-year German textbooks. In J. Levis, R. Mohammad, M. Qian, & Z. Zhou (Eds.) *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching Conference* (pp. 229–242). Ames, IA: Iowa State University.

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

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, *36*, 326–344.

APPENDIX

Emergent Themes from Survey Question about Textbook Approach to Pronunciation Instruction

Category	French	German	Spanish	Total
Sounds	51%	44%	36%	44%
Listen and repeat	27%	31%	21%	26%
Audio/video recordings	16%	21%	24%	20%
Online implementation	13%	5%	33%	18%
Words	11%	18%	14%	14%
Vocabulary connection	11%	8%	12%	10%
Incidentally approached	16%	3%	10%	10%
Recording of student voices	9%	5%	14%	10%
Rule-learning	11%	8%	7%	9%
Cross-linguistic comparisons	7%	3%	12%	7%

Miller, J. (2018). Bridging the gap between pronunciation research and proficiency-based teaching. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 111-119). Ames, IA: Iowa State University.

BRIDGING THE GAP BETWEEN PRONUNCIATION RESEARCH AND PROFICIENCY-BASED TEACHING

Jessica Sertling Miller, UW-Eau Claire

Although many studies have shown that explicit instruction is effective in increasing learners' L2 pronunciation skills (Aliaga-García, 2007; Lord, 2005; Sturm, 2013), lack of phonetic preparation in language education programs and logistical challenges inherent to oral assessments are often cited as possible obstacles to its inclusion (Derwing, 2010; Yates, Zielinski, & Pryor, 2011). To investigate those claims, this study reports on an online survey of 120 instructors of French. Results indicated that instructors value pronunciation, deem to have received adequate training, and feel they draw sufficient attention to it in class. However, activities designed to develop correct pronunciation are seldom offered to learners, highlighting conflicts between the instructors' perceived importance of pronunciation and its relative absence. This paper presents current trends and challenges in pronunciation instruction, and discusses possible ways to bridge research and teaching by taking into consideration the proficiency movement of the last two decades (Brooks & Darhower (2014; Kissau, 2014) as well as teachers' and learners' needs.

Keywords: proficiency, pronunciation, K12, instruction

INTRODUCTION

While there has been an increased interest in investigating the effects of teaching pronunciation in the past decade (Derwing & Munro, 2009; Lord 2005; Trofimovitch, 2013), and despite the fact that explicit instruction has been shown as effective (Aliaga-García, 2007; Lord, 2005; Miller, 2012; Sturm, 2013), there remains a gap between theory and practice as research findings are not routinely put into practice (Gordon, Darcy, & Ewert, 2013). Reasons for setting aside that component of second language acquisition could be, among possible others, lack of teacher training in that area (Derwing, 2010; Foote, Holtby, & Derwing, 2011), challenges related to assessing pronunciation (Yates, Zielinski, & Pryor, 2011), and interferences from current pedagogical practices (Henderson et al., 2012; Kirkova-Naskova et al., 2013). Some scholars have advocated for more communicative components in pronunciation instruction to develop students' intelligibility and comprehensibility (Celce-Murcia, Brinton, & Goodwin, 2010; Hinkel, 2006; Morin, 2007). The widespread proficiency movement of the last two decades (Kissau, 2014), which has shifted the focus toward communicative competence and away from accuracy, may explain why instructors are reluctant to teach pronunciation, a skill that is most effectively taught explicitly, i.e., without apparent communicative authenticity. This investigation examines the current state of pronunciation instruction in introductory French courses to better understand today's pedagogical approaches and what main challenges prevent the inclusion of pronunciation instruction in beginner courses.

BACKGROUND

The proficiency movement

The Proficiency Movement is a ubiquitous language teaching technique in American classrooms today, especially at the K-12 Level (Kissau, 2014). Rather than focusing on grammatical correctness, proponents of this approach concentrate on developing their learners' general proficiency and tolerate inaccuracies as long as communication is successful. This shift could be the consequence of the increased use of tests sanctioned by ACTFL in educational and professional contexts since the 1990s, combined with the popularity of the Communicative Language Teaching Approach (CLT) since the 1980s. Today, 23 states use ACTFL testing for some component of their teacher certification process (NCSSFL). The Oral Proficiency Interview (OPI) in particular is often a source of stress for the candidates who view it as a difficult challenge (Burke, 2013). In a large-scale study, Glisan, Swender and Surface (2013) found that 54.8% of the teacher candidates tested met the required oral proficiency level. If language education programs want to increase their score in an effort to attract students, they would logically focus on proficiency to better prepare their candidates to the test.

Brooks and Darhower (2014) examined the contexts and practices of the three undergraduate language teacher education programs identified by Glisan et al. (2013) as having the highest passing rates. Among common practices that likely contribute to their success, a crucial finding was that most faculty in those programs had been trained through ACTFL and transferred that knowledge to their courses in order to build awareness of ACTFL guidelines in the student body. In turn, student teachers who have been exposed to it will probably design their own classes using similar approaches and extend the Proficiency Movement.

Pronunciation instruction

Given the importance of oral communication within the Proficiency Movement, it seems vital for instructors to emphasize the role of correct pronunciation and to give learners opportunities for oral practice. However, research has shown that speaking causes learners more anxiety than other language skills, especially in front of the entire class (Phillips, 1992). This decreases their Willingness to Communicate (WTC), defined by MacIntyre, Clément, Dörnyei, and Noels (1998, p. 547) for second language learning as a "readiness to enter into discourse at a particular time with a specific person or persons, using a L2." Baran-Łucarz (2014) has shown a correlation between students' WTC and their level of pronunciation anxiety, which often comes from a negative self-perception of their oral skills and a fear of negative judgment, regardless of their proficiency level. The author suggests that pronunciation involves the highest range of emotions in learners, which makes it challenging to teach and to learn.

For instructors to overcome the challenge of engaging students in emotionally-loaded oral practice, one might assume education programs include substantial training in that domain. Research indicates that may not be the case (Foote et al., 2011; Kirkova-Naskova et al., 2013; Murphy, 1997). Foote et al. (2011) posit that lack of training may negatively affect teachers' self-confidence in their ability to tackle pronunciation instruction, possibly discouraging them from trying. Kirkova-Naskova et al. (2013) found that teachers' pronunciation training in Europe is largely phonetic, not pedagogical. They also showed that European instructors generally don't transfer their phonetic training to their classrooms. They hypothesize that this could be due to

lack of pedagogical preparation, assessments not testing pronunciation, lack of time and other resources, or dominance of the CLT approach in the Common European Framework of Reference (CEFR), which favors authentic situations.

Research questions

Several studies indicate the need for a more systematic inclusion of pronunciation instruction in language courses (Derwing, 2010; Kendrick, 1997; Hardison, 2010; Mompean, 2005; Saalfeld, 2011). Explicit lessons have also shown to be effective (Aliaga-García, 2007; Lord, 2005; Miller, 2012, Sturm, 2013), but proficiency-based approaches value authentic communicative tasks over explicit instruction, which may explain why pronunciation instruction has long been neglected (Morin, 2007; Saalfeld, 2011). In order to reconcile research and practice, this study seeks to gain knowledge on the current state of pronunciation instruction in introductory French courses. This broad interrogation is broken down into smaller questions:

1. What teaching strategies are currently employed to teach pronunciation?
2. What impact does their training have on pronunciation instruction?
3. How do teachers view pronunciation instruction compared to other skills?
4. What challenges do they currently face?

METHODS

Procedure and participants

The author-researcher created an anonymous electronic questionnaire with Qualtrics and distributed to a North-American online network of teachers of French. Completion required approximately eight minutes. Each section included both multiple-choice and open-ended questions. A total of 162 respondents initiated the online questionnaire and 120 completed it fully. A vast majority (90%) were K-12 instructors, with most (61%) being high school teachers. Among them 43% have been teaching 10 years or less.

Analyses

Frequency statistics were used to analyze the quantitative data collected through the survey. The qualitative information was coded following a grounded theory approach (Glaser, 1992). As responses were examined in a Word document produced by Qualtrics, emerging recurrent themes were tagged with unique codes. The Find feature of Word helped quantify those occurrences and turn them into percentages.

RESULTS

Teaching strategies

When asked if they teach or draw students' attention to pronunciation, almost all (97%) answered positively. Among a choice of four explicit teaching strategies, the most commonly selected was explaining spelling-to-sound correspondences (90%), followed by providing articulatory tips (78%). Given a choice of four implicit teaching strategies, a majority opted for modeling pronunciation and having learners repeat after them (94%), followed with giving immediate corrective feedback via recasting (82%). Encouraging learners to figure out pronunciation rules on their own was also a popular choice (44%). Approaches that hadn’t been proposed earlier

emerged in open comments: focusing on sounds that can create communication difficulties, using English letters and sounds to help learners understand or memorize French pronunciation, as well as using the Accelerative Integrated Method hand signs.

Assessment

Three fourths (74%) of participants reported assessing learners' pronunciation. The types of assessments used vary widely. Most seemed to consist of oral interviews (68%) and in-class presentations, either in groups (62%) or individually (51%). Pronunciation also seemed to often be assessed during regular class activities (43%). But assessing does not automatically mean grading. One third of respondents (30%) reported not grading their learners' pronunciation. When they do grade pronunciation in an introductory French course, most instructors indicated taking off points only when mispronunciation interfered with the speaker's intended message: "It's always a component of comprehensibility. If you can't be understood, it needs improvement."

Training

Half the participants (51%) reported having received adequate or thorough preparation to help them instruct pronunciation. A few (17%) reported having received none. According to respondents, their preparation either had a positive impact (44%) or no impact (42%) on the way they teach pronunciation in introductory French, with 5% reporting it negatively affected it. Half the respondents (49%) explained that their main exposure to French pronunciation instruction as students was typically through a single corrective pronunciation course as part of their French program's curriculum. The next most frequent answer was that they received no training (20%). A few (17%) wrote that pronunciation was discussed minimally in their methods courses. The rest mentioned professional development such as workshops or mentoring as helpful (14%). One person wrote: "I received training in pronunciation, but not how to teach pronunciation. After learning it myself, though, I realized how much it would have helped me to know earlier on. Therefore, it is a goal to eventually incorporate it in my curriculum so that my students can benefit from it. Just not quite there yet." A few noted that their teacher training courses deemphasized the importance of pronunciation, which is why they do not focus on it in their classes: "[…] we were often told by many professors not to focus too hard on pronunciation, unless it was a big enough problem where the words were either incomprehensible or the meaning was changed when mispronounced. This is in order to create a welcoming and open learning environment where students feel safe to take language risks".

Relative importance

Participants were asked to rate how important it is to them to teach pronunciation in relation to other language skills on a scale from least to most important (1 to 5). Results showed that pronunciation is on average regarded as somewhat more important (2.96) to teach as other skills. When asked to rate how important it is for their students to acquire basic pronunciation skills in relation to other skills, the number increased slightly. Answers indicated that it is on average regarded as a bit more important (3.07) to learn than other skills. A quarter of those who explained their answers in the open comments section (27%) believe that if pronunciation is not taught in introductory courses, fossilization occurs and creates lasting communication problems. But just as many respondents believe that exposing beginner students to oral French without explicit pronunciation tips, i.e., acquisition by osmosis, is sufficient: "As in any language learned naturally, my students will learn to pronounce the words through repetition and with time".

Challenges

Substantial obstacles (Table 16) are lack of time (16%), fear to lower the learners' WTC (10%), difficulty balancing all language skills (9%), difficulty assessing pronunciation (7%), and perception that instruction is only effective on motivated learners (3%). One respondent wrote: "I think it is incredibly important to teach pronunciation from the beginning so learners develop all skills at an equal pace, rather than having to "correct" their pronunciation down the road. As with everything else, the challenge is dividing class time to address all aspects of communicative competence. An additional challenge is addressing pronunciation adequately without causing frustration or boredom."

DISCUSSION

Effects of proficiency-oriented pedagogies

Pronunciation instruction is important to teachers but challenges seem to prevent them from including it systematically in active learning tasks, based on the survey responses and a parallel observation of 30 hours of French instruction in area schools (not reported in this paper but presented at PSLLT 2017). Instead, pronunciation instruction seems to be included reactively, having found its place in corrective feedback and follow-up comments. The role of pronunciation instruction has shifted as demands and constraints evolved in second language teaching and learning. Most participants indicated valuing pronunciation instruction and wishing to incorporate more in class, but being unable to for reasons beyond their control: proficiency tests, college placement exams, and imposed curricula reward the development of vocabulary and grammar. ACTFL guidelines rely on patterns of communication breakdowns and successes to evaluate a speaker's proficiency level. At the novice and intermediate levels, a speaker only needs to be understood by sympathetic listeners accustomed to dealing with non-native speakers. In beginner courses, instructors are likely to understand their learners despite pronunciation inaccuracies because of their experience and training. There is therefore little incentive to work specifically on pronunciation at that stage lest students might be discouraged and drop French. At a time when French programs have become electives and suffer from low enrollment numbers, encouraging students to enroll in their courses is a definite concern for French teachers. The Proficiency Movement emphasizes a natural acquisition of those skills through games, information-gap tasks, and role-plays reflecting authentic situations and cultures. Drills and fill-in-the blanks exercises may be used as support, but are typically not central to that approach. Developing proficiency-oriented material to boost pronunciation accuracy is possible but time-consuming, and has typically not been modeled and therefore teachers would need to create it from scratch.

Effects of teacher training

Foote et al. (2011) hypothesized that instructors might lack self-confidence due to inadequate preparation. In this study, a few participants cited self-confidence as a challenge, but most reported having received adequate training. However, the training to which they referred was generally phonetic, not pedagogical, paralleling European findings (Kirkova-Naskova et al., 2013). As noted by some respondents, pronunciation instruction is typically either barely approached in methods courses, if at all, or it is discouraged because its emotional risks outweigh its communicative benefits, as discussed by Baran-Łucarz (2014). Despite the demonstrated

effectiveness of explicit instruction, it remains overlooked in methods courses, possibly because it conflicts with proficiency-based training favoring implicit approaches. That absence of pronunciation in methods courses might send the signal that pronunciation is of little importance.

Learners' needs

This study demonstrates that instructors of French want to infuse more pronunciation instruction in their introductory courses. Some indicated students even request it, echoing previous research findings (Derwing & Munro, 2009; Saalfeld, 2011). While explicit pronunciation instruction could be effective within instructors' time constraints, it seems to go against instructors' pedagogical beliefs; on the other hand, acquisition by osmosis is in accord with their pedagogical practices but is not effective within their time constraints. The most satisfying compromise currently is to reserve pronunciation instruction to reactive remarks, i.e., when explicit clarifications can be rationalized as acceptable. This is supported by numerous comments describing instructors' preferred strategy as seeking opportunities to give learners feedback as the needs arise.

Classroom implications

Researchers, instructors, and learners seem to be in agreement that pronunciation instruction should receive more attention. The types of activities promoting the development of pronunciation skills should (1) make pedagogical sense to today's teachers, (2) be ready to use, support national standards, and comply with local school missions, (3) be effective as demonstrated by research, (4) encourage learners to enroll in French courses. Tasks anchored in authentic and current situations relevant to the learners should be part of the pronunciation arsenal, as well as games that present entertaining ways to elicit and repeat language forms that might not emerge naturally, similar to the function of OPI's role-plays. Since pronunciation instruction may raise negative emotions in learners, teachers understandably may see its inclusion as risky, as one respondent remarked: "Of course I want to encourage my students to the best pronunciation possible, but a teacher [must] respect that fragile balance point between effective communication and self-confidence."

Future directions

What elements would therefore make pronunciation-focused activities more likely to be implemented by proponents of proficiency-based teaching in introductory French courses, and demonstrate effectiveness on learners' proficiency levels? Morin (2007) explored the possibilities for Spanish, but research is needed to identify pronunciation features linked to certain functions of proficiency levels. Once key characteristics have been identified, researchers need to provide concrete teaching ideas and ready-to-use material that will fit the teaching styles of local K12 teachers and match their students' needs.

CONCLUSION

Pronunciation instruction has not exactly been neglected in introductory French courses: while it currently seems to rarely be integrated proactively, it remains included reactively in teachers' immediate corrective feedback and follow-up comments. Proficiency-oriented pedagogy combined with constraints beyond instructors' control appear to have contributed to this shift. More research is needed to identify specific pronunciation features correlating to proficiency functions, especially as framed by ACTFL since their standards and oral tests are widely used by

teachers and learners. Investigating pronunciation error types and categorizing them into levels would also help target what phonological items are worthy of inclusion in beginner courses. For example, which errors hinder communication most (Novice), which do sympathetic listeners find acceptable (Intermediate), which are acceptable to listeners not accustomed to dealing with non-native speakers (Advanced), and which have the least impact (Superior).

ABOUT THE AUTHOR

Jessica Sertling Miller (Ph.D., University of Illinois at Urbana-Champaign), a native of Strasbourg, France, is a Professor of French at the University of Wisconsin-Eau Claire. She teaches all levels of undergraduate courses and supervises students in the French Education program. Her recent research has focused on proficiency-based teaching and learning.

Phone: 715-836-4267

Email: millerjs@uwec.edu

REFERENCES

Aliaga-García, C. (2007). The role of phonetic training in L2 speech learning. *Phonetics Teaching & Learning Conference*. Retrieved October 31, 2017, from http://www.phon.ucl.ac.uk/ptlc/proceedings/ptlcpaper_32e.pdf

Baran-Łucarz, M. (2014). The link between pronunciation anxiety and willingness to communicate in the foreign-language classroom: the Polish EFL context. *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, 70(4), 445-473.

Brooks, F. & Darhower, M. A. (2014). It takes a department! A study of the culture of proficiency in three successful foreign language teacher education programs. *Foreign Language Annals*, 47(4), 592-613.

Burke, B. (2013). Looking into a crystal ball: Is requiring high-stakes language proficiency tests really going to improve world language education? *The Modern Language Journal*, 97(2), 531-534.

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

Derwing, T. M., Munro, M. J., (2009). Putting accent in its place: rethinking obstacles to communication. *Language Teaching*, 42(4), 476-490.

Derwing, T. M. (2010). Utopian goals for pronunciation teaching. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, (24-37). Ames, IA: Iowa State University. Retrieved October 31, 2017, from https://apling.engl.iastate.edu/alt-content/uploads/2015/05/PSLLT_1st_Proceedings_2009.pdf

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

Glaser, B. G. (1992). *Basics of grounded theory analysis: Emergence vs. forcing*. Mill Valley, CA: Sociology Press.

Glisan, E., Swender, E., & Surface, E. (2013). Oral proficiency standards and foreign language teacher candidates: Current findings and future research directions. *Foreign Language Annals*, 46(2), 264-289.

Gordon, J., Darcy, I., & Ewert, D. (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, (194-206). Ames, IA: Iowa State University. Retrieved October 31, 2017, from http://jlevis.public.iastate.edu/pslltconference/4th%20Proceedings/Gordon_Darcy_Ewert%20PSLLT%202012.pdf

Hardison, D. M. (2010). Trends in teaching pronunciation. *CLEAR Newsletter (Center for Language Education and Research, Michigan State University),* 14(2), 1-5.

Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Cunningham, U., & Curnick, L. (2012). The English pronunciation teaching in Europe survey: Selected results. *Research in Language*, 10(1), 5-27.

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40(1), 109-131.

Kendrick, H. (1997). Keep them talking! A project for improving students' L2 pronunciation. *System*, *25*(4), 545-560.

Kirkova-Naskova, A., Tergujeff, E., Frost, D., Henderson, A., Kautzsch, A., Levey, D., Murphy, D., & Waniek-Klimczak, E. (2013). Teachers' views on their professional training and assessment practices: Selected results from the English Pronunciation Teaching in Europe survey. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, (29-42). Ames, IA: Iowa State University. Retrieved October 31, 2017, from http://jlevis.public.iastate.edu/pslltconference/4th%20Proceedings/Kirkova-Naskova%20et%20al%20%20PSLLT%202012.pdf

Kissau, S. (2014). The impact of the Oral Proficiency Interview on one foreign language teacher education program. *Foreign Language Annals*, Vol. 47(3), 527-545.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 88, 557-567.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, 82(4), 545-562.

Miller, J. S. (2012). Teaching French pronunciation with phonetics in a college-level beginner French course". *The NECTFL Review*, 69, 47-68.

Mompean, J. A. (2005). Taking advantage of phonetic symbols in the foreign language classroom. *Phonetics Teaching & Learning Conference*. Retrieved June 15, 2015, from https://www.phon.ucl.ac.uk/home/johnm/ptlc2005/pdf/ptlcp28.pdf

Morin, R. (2007). A neglected aspect of the standards: preparing foreign language Spanish teachers to teach pronunciation. *Foreign Language Annals*, 40(2), 342-360.

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

Phillips, E. M. (1992). The effects of language anxiety on student's oral test performance and attitudes. *The Modern Language Journal*, 76(1), 14-26.

Saalfeld, A. K. (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*, (144-152). Ames, IA: Iowa State University. Retrieved October 31, 2017, from http://jlevis.public.iastate.edu/2010%20Proceedings%2010-25-11%20-%20B.pdf

Sturm, J. L. (2013) Explicit phonetics instruction in L2 French: A global analysis of improvement. *System*, 41(3), 654-662.

Trofimovich, P. (2013). Interactive alignment: Implications for the teaching and learning of second language pronunciation. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, (1-9). Ames, IA: Iowa State University. Retrieved October 31, 2017, from http://jlevis.public.iastate.edu/pslltconference/4th%20Proceedings/Trofimovich%202012%20(1).pdf

Yates, L., Zielinski, B. & Pryor, E. (2011). The assessment of pronunciation and the new IELTS Pronunciation scale. IELTS Research Reports. Retrieved June 15, 2015, from http://www.ielts.org/pdf/Vol12_Report1.pdf

Uzun, T. & Ay, S. (2018). Preservice English teachers' perspectives on pronunciation. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 120-128). Ames, IA: Iowa State University.

PRESERVICE ENGLISH TEACHERS' PERSPECTIVES ON PRONUNCIATION

Tarık Uzun, Ankara Yıldırım Beyazıt University

Sıla Ay, Ankara University

The goal of this study is to investigate pre-service English teachers' views, needs and preferences about learning and teaching pronunciation. Using mixed-method approach, data have been collected through a questionnaire (N=147) and additional face-to-face interviews (N=5) with prospective English teachers enrolled in an English Language Teaching (ELT) department at a state university in Turkey. A 40-item questionnaire (adapted from Sardegna & Kusey, 2014 and Seyedabadi, Fatemi & Pishgadam, 2014 with additional items prepared by the researchers) was used to provide insights into pre-service English teachers' views, preferences and goals about learning and teaching pronunciation. In addition, face-to-face interviews were carried out with students who have prior voluntary or semi-professional teaching experience. Findings of the study reveal that pre-service English teachers consider learning pronunciation as a key area in learning a foreign language, however, they are not prepared to teach it with informed techniques in a classroom setting.

INTRODUCTION

Investigating teachers' views, beliefs, awareness levels and knowledge about various aspects of teaching practice has been a popular line of enquiry. Studies that put the spotlight on teachers' cognitive processes fall under the title of teacher cognition research. Borg (2003) defines teacher cognition as *"the unobservable cognitive dimension of teaching – what teachers know, believe, and think*" and bases the assumptions of the concept on the following argument: *"teachers are active, thinking decision-makers who make instructional choices by drawing on complex, practically-oriented, personalised, and context-sensitive networks of knowledge, thoughts, and beliefs."* (pp. 81). Along with increasing the expertise and awareness of the teachers in profession, teacher cognition research can also affect the professional development of prospective teachers (Baker & Murphy, 2011).

It is evident that there is an increased interest in researching teacher cognitions in pronunciation, however, as also mentioned by Baker & Murphy (2011), less attention has been paid to this field within the domain of pronunciation teaching so far. Studies focusing on preservice English teachers (from here onward PrET) and their developing cognitions about pronunciation teaching are even more scarce. Baker (2011) underlines the essential role of teacher education in shaping teachers' cognitions in pronunciation pedagogy, and claims that, without such an education, their knowledge of pronunciation pedagogy along with their confidence in teaching pronunciation would be limited. Similarly, Burri (2015) and Couper (2016) indicate that teacher education programs and teacher educators need to focus on pronunciation teaching techniques, which could increase teachers' or PrETs' confidence and help develop their cognitions.

A number of studies that were carried out in different countries found that teachers feel unprepared to teach pronunciation because of insufficient training they received in teacher

education programs. For instance, Henderson et al. (2015) conducted a survey in seven countries across Europe, where most of the participants appeared to be amateurs in pronunciation teaching since they received no or little training in how to teach it. Fallang (2016) mentions a similar finding that participants did not receive extensive teacher training regarding pronunciation in Norway. In a study in Canada, Breitkreutz, Derwing and Rossiter (2001) found out that it was necessary to integrate pronunciation in a communicative classroom while teachers lacked the knowledge to do so, which is associated with the lack of sufficient training and training opportunities. 10 years after this study, Foote, Holtby and Derwing (2011) aimed at reexamining the state of teaching pronunciation in ESL classes in Canada. Having scanned the websites of Canadian universities that offer TESL programs, the researchers could find only 6 universities where pronunciation-specific courses are offered.

At this point, the state of pronunciation teaching in teacher education programs in Turkey needs to be explored extensively. In the country, graduates of English Language Teaching (ELT) departments are officially entitled as *English teachers.* They are allowed to work at state schools if they meet a number of other official requirements as well, such as taking a central examination. In addition, graduates of other departments like English Language and Literature or Linguistics can also teach English at state schools if they obtain a teacher training certificate and meet other relevant criteria. Higher education institutions, on the other hand, follow different procedures while employing English teachers.

Most of the ELT departments in Turkey use a dedicated course program designed by Turkish Higher Education Council (HEC). In this 4-year program, a number of linguistic and pedagogical courses are offered to PrETs. In terms of pronunciation, two pronunciation courses can be found in the program: Listening and Pronunciation I and II. HEC also describes possible content for each course, however, basically, faculty staff members are free to decide on what to teach and revise the content in these classes.

PrETs' perspectives related to pronunciation teaching in ELT programs have not been widely investigated in Turkey. As a preliminary study of a larger research project to be carried in a number of state universities, the present study aims to explore pre-service English teachers' views, needs and preferences about learning and teaching pronunciation.

METHOD

The study was conducted at the English Language Teaching Department of a Turkish state university in the Spring semester of 2017. The participants were all students enrolled in the department at the time of data collection.

Data were collected in Turkish using mixed-method approach. In the first step, PrETs' views, preferences and goals about learning and teaching pronunciation were investigated with the help of a questionnaire (n=147). Secondly, semi-structured-interviews were carried out with students from the same department (n=5) with the purpose of gaining broader insights into PrETs' perspectives about learning and teaching pronunciation. Before both steps, participants were asked to complete a demographic questionnaire including basic information like gender, age, and class.

The Questionnaire

In order to find out about PrETs' perspectives in learning and teaching pronunciation, a

questionnaire with 40 likert type items (1-5; 5 being "I completely agree" and 1 "I completely disagree") was designed. Out of these 40 items, 26 of them were adapted from two existing scales (13 items from Sardegna & Kusey, 2014 and 13 others from Seyedabadi, Fatemi & Pishgadam, 2014) and the remaining 14 were developed by the researchers. During the selection and item-writing processes, it was aimed to address PrETs' general views on learning and teaching pronunciation, perspectives in learning and teaching English pronunciation and personal goals, expectations and needs.

The questionnaire was prepared in Turkish in order to be fully understandable to the target population. The items adapted from the existing scales were translated into Turkish and the rest of them were all prepared in Turkish. To assure reliability in translation, the selected items were translated separately by the two researchers and then compared for the final form.

The questionnaire was presented to 147 PrETs who speak Turkish as their first language. Along with international students who have no knowledge of Turkish, those who speak Turkish at a low level were not included in the sample either.

The internal consistency of the instrument was investigated through Cronbach's Alpha and the Alpha coefficient value was found 0.716. The value equal to or above 0.7 is considered 'acceptable' (George & Malley, 2003), therefore, the questionnaire was found reliable for descriptive analyses.

Semi-Structured Interviews

Semi-structured face-to-face interviews were also conducted at the same university (n=5). Interviewees, were third year students with previous teaching experience (private tutoring, voluntary teaching or institutional experience). It was considered that recruiting participants with teaching experience could add depth to the study results.

Interviews were carried out in Turkish and all the participants were Turkish native speakers. Each interview lasted around 8-12 minutes and recorded with a voice recorder. 10 questions (see Appendix) with the following three themes were asked to the participants:

1. Preferences, goals, awareness levels, and self-efficacy beliefs related to their own pronunciation
2. Views on the role of pronunciation in teacher education
3. Views on teaching pronunciation

The interview data were transcribed, coded and analyzed for recurring themes. In order to achieve inter-coder reliability, the two researchers coded the transcriptions separately and findings were compared. While most of the themes were found similar, a consensus was still sought and reached on the divergent items.

RESULTS

THE QUESTIONNAIRE

Demographics of the Participants

Almost all of the participants who completed the questionnaire were found to be between the ages of 17-25 and the number of female participants was higher than male students. Besides

these, numbers of participants for first to third year students were close to one another while the number of 4th year students was considerably lower compared to the others. Details of participant demographics can be found in tables 1-3 below:

Table 1

Age Groups

Age	N	%
17-25	145	98.6
26-35	2	1.4

Table 2

Gender

Gender	N	%
Female	91	61.9
Male	56	38.1

Table 3

Years (Classes)

Classes	N	%
1	58	39.5
2	34	23.1
3	50	24
4	5	3.4

Questionnaire Findings

Participant responses have been analyzed descriptively with mean scores, standard deviations, percentages and frequencies. On the whole, the results indicate a strong tendency to rate the items positively in general. To be more specific, the mean scores of 20 items were above 4 while only 3 items were rated below 3. More specifically, the findings will be reported in detail below.

Agreement on the Importance of Pronunciation

According to the results of the survey, PrETs believe that learning and teaching pronunciation is essential for both learners and teachers. For instance, more than 90% of the participants rated each of the following items 4 or 5: "Pronunciation is the inseparable part of English language learning.", "I would like to have good pronunciation", "People who want to teach English must have good pronunciation." and "In my opinion, a teacher must definitely spend time on

pronunciation in his/her class". This finding shows a significant tendency to consider pronunciation a key element in learning and teaching English. However, the item "Learning to pronounce well is the most important part of learning a language" received more diverse reactions than the others shown above. Although the majority of the participants still agreed with it, 46 participants (31.3%) remained undecided while 20 (13.6%) others disagreed. This implies that PrETs value the role of pronunciation, yet they still think it is not the most important element of the language.

Pronunciation and Other Language Skills

The results of the questionnaire also suggest that participants see connections between pronunciation and reading, listening and speaking. The items which underline a relationship between pronunciation and each of the above-mentioned skills were 'agreed' by at least 80%. It even reaches 90% for the item "Pronunciation has a mutual effect on listening skill."

Pronunciation and Affective Dimensions

The results also reveal a strong connection between affective factors and pronunciation. The majority of the participants acknowledge the role of pronunciation in reducing language learner's anxiety (122 participants; 83%) and raising English language learners' degree of self-confidence (143 participants; 98%). Similarly, 129 PrETs (87.7%) agree that "An English language learner's poor pronunciation influences his/her willingness to communicate." Despite these results, participants seem to get less confident when it comes to their own pronunciation. 32% of them were undecided and 5.4% disagreed with the following item: "I feel confident that my pronunciation will be understood when I talk."

Readiness to Teach Pronunciation

As future English teachers, participants of this study did not seem to be prepared to teach pronunciation with informed techniques. One of the lowest mean-scored items, "I know how to teach pronunciation to my students in the future.", clearly refers to the problem of lack of sufficient training. While 52 students (35.4%) disagree with the statement, 61 (41,5%) picked the "I'm undecided" option.

Participants were also asked whether pronunciation needs to be taught explicitly in class or it is a skill that should be acquired more implicitly. Analyses reveal that 65 (44,2%) participants were undecided about teaching pronunciation explicitly. On the other hand, 69 participants (47%) disagreed with an implicit way of learning while 51 others (34,7%) were undecided. These two items imply that PrETs are not sure about their preferred approach in learning and teaching pronunciation.

Two other items complete the argument with the preferences of participants. One points out to their need for a 'how to teach pronunciation' course in their undergraduate course program and 142 participants (95.2%) supported this idea. Also, 78.9% of the participants were of the opinion that they would be willing to take an elective pronunciation course if it were offered in their undergraduate program.

Native-like vs. Intelligible Pronunciation

The results of the questionnaire demonstrate that participants tend to set native-like proficiency as an attainable goal in learning English. To illustrate, 126 participants (85.7%) believe that "The aim of learning English pronunciation is to achieve a native-like accent." and 134 of them

(91.1%) personally aim to reach a native-like accent. The item "Intelligible pronunciation will be adequate for the learners of English.", on the other hand, created uncertainty for many participants since 59 participants (40.1%) were undecided while 54 of them (38.1%) agreed with it.

Accent Preferences

Participants were asked about their preferences of British or American English in two separate items. The results seemed quite similar to each other without referring to any strong tendency towards one. Similarly, 50 participants (34%) remained undecided when asked about which accent they actually speak in English.

THE INTERVIEWS

The 10 interview questions were grouped into three themes; therefore, the interview data were analyzed basically in three clusters. In order to reach the most commonly expressed views, the ones mentioned by all or at least the majority of the interviewees were taken into consideration and the rest was discarded. Below can be found the basic categories and concepts falling under each theme:

Group 1. Preferences, goals awareness levels, and self-efficacy beliefs related to their pronunciation

PrETs generally indicated a feeling of incompetency in pronunciation. When their preferred ways to improve their pronunciation was enquired, most of them referred to watching TV broadcasts. In relation to this finding, they demonstrated a significant preference to speak in American English as they are exposed to it in their favorite TV series and movies.

It has also been found out that participants consider linguistic differences between the two languages as a possible cause for intelligibility problems for Turkish speakers. In this regard, differences in segmentals and manner of articulation were frequently addressed by the participants.

Group 1. Views on the role of pronunciation in teacher education

Interviewees generally pointed out that their departmental courses focused on issues like articulation of individual sounds and International Phonetic Alphabet (IPA). However, according to the participants, these courses lacked the possible techniques to teach pronunciation in actual classroom settings. They all underlined their expectations to learn practical teaching techniques or tips and demanded more extensive integration of listening and speaking skills in their classes.

Group 3. Views on teaching pronunciation

Participants believe that nonnative teachers are capable of teaching pronunciation and being a native speaker cannot be a prerequisite to teach it. This widespread belief among the PrETs was commonly explained through positive experiences with their own nonnative teachers in the past.

PrETs all acknowledge the importance of focusing on pronunciation in class and being a good model for their students. They frequently pointed out that teachers need to be perfect in pronunciation. Similarly, they also expressed their desire to achieve a native-like accent.

In connection with the findings in Group 2, these interviewees had difficulty proposing suitable

techniques or ways they would use to teach pronunciation. Their suggestions were quite limited lacking variety. These results were also found in line with questionnaire findings.

DISCUSSION

The goal of this study was to explore PrETs' views, needs and preferences on learning and teaching pronunciation. In Turkey, as nonnative English speakers, PrETs need to improve their own pronunciation skills in order to be good speech models for their students. Their future roles as teachers, however, will require finding ways to bring pronunciation into their classes to help other learners, possibly at different age groups. Without a doubt, it is a matter of decision for teachers or program coordinators whether or not to focus on pronunciation in their classes. As MacDonald (2002) puts it, teachers could be reluctant or feel lack of confidence in teaching pronunciation but developing their skills in integrating pronunciation could be one of the helpful solutions to overcome this challenge.

The results of this present study can be summarized in two broad categories.

In terms of learning English, PrETs tended to consider intelligible pronunciation an ideal goal for English learners (instead of aiming for a native-like accent) but the common argument was that native speaker level pronunciation was essential for teachers as they frequently stated "being a teacher requires perfection". This tendency also implies that they are highly motivated to keep improving their skills in pronunciation.

In terms of teaching English, PrETs considered pronunciation as a key area, however, they did not seem to be prepared to teach pronunciation with informed techniques in different classroom settings. When asked about what techniques they would use in pronunciation teaching, they could only suggest intuitive ideas, lacking variety. This finding was found in line with other studies in which teachers in different contexts referred to a lack of or insufficient training on pronunciation teaching in their teacher education programs (Breitkreutz, Derwing & Rossiter, 2001; Foote, Holtby, Derwing, 2011; Henderson et al, 2015).

Our observation is that pronunciation is usually neglected in teacher education in Turkey as in many other countries; however, further investigation is necessary to figure out what the case is in different institutions in higher education. Such attempts will be helpful in determining the priorities in shaping the future of pronunciation teaching and research in the country.

ABOUT THE AUTHORS

Tarık Uzun has been working as an Instructor of English and Turkish since 2005. At Ankara Yıldırım Beyazıt University School of Foreign Languages, where he is currently based, he coordinated the Independent Learning Center (ILC) for 5 years. He is also a PhD candidate at Ankara University Department of Linguistics. In 2017-2018 academic year, he visited Iowa State University Department of Applied Linguistics and Technology as a Fulbright Visiting Research Scholar to work on his doctoral research. His academic interests include pronunciation research, learner autonomy and self-access language learning.

Contact Information

E-mail: uzuntarik@yahoo.com

Address: Ankara Yıldırım Beyazıt Üniversitesi Yabancı Diller Yuksekokulu Ayvalı Mah. Gazze Cad. No: 7 Keçiören, Ankara Türkiye

Tel: +905062119889

Sıla Ay has been a lecturer in Ankara University, Foreign Language Teaching Division since 1995. She had her PhD on foreign language reading skills. In the last five years, she has focused mainly on individual differences and foreign language pronunciation. Currently, she is working as an associate professor in the Applied Linguistics Department of Ankara University. She has trained many language teachers who are currently working all around Turkey in different schools and universities.

Contact Information

E-mail: silaay@gmail.com

Address: Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi 06100 Sıhhiye, Ankara Türkiye

Tel: +905325577800

REFERENCES

Baker, Amanda A. (2011). Pronunciation pedagogy: Second language teacher cognition and practice [Doctoral Dissertation]. Georgia State University, Atlanta, USA. Retrieved from http://scholarworks.gsu.edu/alesl_diss/16.

Baker, A. A. & Murphy, J. (2011). Knowledge base of pronunciation teaching: Staking out the territory. TESL Canada Journal, 28(2), 29-50.

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

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

Burri, M. (2015). Student teachers' cognition about L2 pronunciation instruction: A case study. *Australian Journal of Teacher Education, 40*(10), 66-87.

Couper, G. (2016). Teacher cognition of pronunciation teaching amongst English language teachers in Uruguay. *Journal of Second Language Pronunciation, 2*(1), 29-55.

Fallang, I. (2016). Teachers' knowledge and practices of pronunciation in lower secondary school in Norway [Masters Dissertation]. The Arctic University of Norway, Tromso, Norway. Retrieved from https://munin.uit.no/handle/10037/9532.

Foote, J. A., Holtby, A. K, Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. TESL Canada Journal, *29*(1), 1-22.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th Ed.). Boston: Allyn & Bacon.

Henderson, A., Frost, D., Tergujeff, E., Kautzsch, A., Murphy, D., Kirkova-Naskova, A., Waniek-Klimczak, E., Levey, D., Cunningham, U. & Curnick, L. (2012). The English pronunciation teaching in Europe survey: Factors inside and outside the classroom:

Selected results. *Research in Language, 10*(1), 5-27.

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

Sardegna, V. G., Lee, J. & Kusey, C. (2014). Development and validation of the learner attitudes and motivations for pronunciation (LAMP) inventory. *System, 47*, 162-175.

Seyedabadi, S., Fatemi, A. H. & Pishghadam, R. (2014). Construction and validation of EFL learners' attitudes toward English pronunciation (LATEP): A structural equation modeling approach. *Theory and Practice in Language Studies, 4*(9), 1929-1940.

APPENDIX

Interview Questions (Translated from Turkish)

1. Do you make conscious efforts to improve your pronunciation? If your answer is 'yes', please explain how.
2. Do you find yourself competent at pronunciation? Can you specify any areas of difficulty for you?
3. Do you speak English with a particular accent/variety (like British, American, Australian, Canadian)? If your answer is 'yes' please explain why you speak in that particular accent.
4. Do you aim to speak English with a native speaker's level of proficiency? Or do you think 'intelligibility' is a more suitable goal for you?
5. What are the factors that could reduce the intelligibility of a learner of English as a foreign/second language?
6. Are you satisfied with the pronunciation that was taught in your undergraduate classes?
7. What should be taught in a pronunciation course in English Language Teaching programs?
8. Do you think that a nonnative English teacher can teach pronunciation? Should pronunciation be taught only by native speakers?
9. Do you think that English teachers should spare time for pronunciation teaching in their class hours? If your answer is 'yes', please explain how much time should be spent for it.
10. As a PrET, do you know any techniques or ways to teach pronunciation to your students? If your answer is 'yes', please give two or three examples.

Kruger, F. (2018). Processing indexical and dialectal variation in a second language. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 129-140). Ames, IA: Iowa State University.

PROCESSING INDEXICAL AND DIALECTAL VARIATION IN A SECOND LANGUAGE

Franziska Kruger, Indiana University

To improve second-language (L2) learners' listening skills, it has been suggested to enhance classroom-input with variation (e.g. more speaker voices, different dialects) to simulate a realistic linguistic landscape in the classroom. However, previous studies reported that L2-learners struggle to distinguish dialects and voices, even at advanced levels. This study explored L2-classroom learners' ability to group words based on indexical variation (speaker voice and dialect). Twenty-seven learners of German (14 beginners, 13 intermediates) and five native speakers classified thirty tokens based on their perceived similarity of voice and dialect in a free classification task. All participants distinguished the stimuli to some degree, but classification accuracy for beginners was significantly less accurate than for native speakers. Intermediates presented with large variation, but accuracy did not differ significantly from native speakers. All groups relied on the same acoustic cues, but their perceptual spaces reveal that NS were more successful than both learner groups at using those cues to differentiate tokens. The findings suggest that L2-classroom learners process indexical variation less efficiently and that word-familiarity might influence their performance. Increasing input variability in classrooms without taking these observations into account could make listening tasks more difficult and hinder learning- and listening-skill development.

INTRODUCTION

The perceptual system for speech processing adapts very efficiently to speaking style, rates, and dialects (Pisoni, 1997), and studies on language variation find that listeners pay close attention to indexical information that helps them interpret the variation inherent in the speech stream (Clopper & Bradlow, 2009). *Indexical information* is speaker-specific information about their region of origin (dialectal variation), as well as idiosyncrasies (e.g. gender, voice quality, Labov, 1972). It is linked to linguistic information and facilitates word recognition and processing even in difficult listening conditions (Cooke, García Lecumberri, & Barker, 2008).

The perceptual system of second-language (L2) learners is reportedly less efficient and less flexible at processing linguistic—e.g. phonetic-phonological—information (Baker & Trofimovich, 2005), as well as indexical information in the L2 (Clopper & Bradlow, 2009; Sullivan & Schlichting, 2000). To date, few studies have explored how L2 classroom learners process indexical information and whether idiosyncratic or dialectal variation is salient enough to be noticed and interpreted accurately during the initial stages of the L2-acquisition process. This study explores L2-German learners' ability to classify words, based on two specific indexical properties (speaker voice and dialect) in relation to their proficiency level, as well as their linguistic background.

LITERATURE REVIEW

The L1-perceptual system is highly attuned to linguistic information, and previous research has shown a strong familiarity effect for standard and dialectal varieties (Clopper & Pisoni, 2004) as well as for voice identification (Goggin, Thompson, Strube, & Simental, 1991). Listeners with exposure to various dialects perform more accurately on categorization tasks than dialect-naïve listeners, i.e. those who have lived in one only region (Pisoni & Clopper, 2004). Nevertheless, naïve listeners do perceive acoustic-phonetic features relevant to distinguish regional varieties and perform above chance levels on categorization tasks (Clopper, Levi, & Pisoni, 2006; Clopper & Pisoni, 2004). For L2 learners, research focuses mainly on their acquisition of linguistic properties, often confirming that learners struggle to acquire the L2 phonological system (Baker & Trofimovich, 2005). These difficulties are mirrored in learners' performance when processing indexical information in their L2 (Eisenstein, 1982; Eisenstein & Verdi, 1985; Köster & Schiller, 1997; Köster, Schiller, & Künzel, 1995).

Clopper and Bradlow (2009) found that native speakers (NS) of American English (AE) were more accurate at sorting talkers into US regions than non-native listeners from various L1 backgrounds. Their L2 participants had been in the U.S. less than 12 months, which may not have sufficed to improve their processing skills for L2-indexical information. Eisenstein (1982) and Eisenstein and Verdi (1985) found varying degrees of accuracy for L2 learners. Beginning and intermediate learners with short and longer length of residence (LOR) in New York were less accurate than NSs when asked to distinguish AE dialects, but advanced learners' performance did not differ from the NSs. It remains unclear if proficiency or amount of input influenced learner performance; the authors provided little information on learner LOR and coded proficiency levels based on course placement at colleges. While these studies indicate that learners in immersion settings can improve, no studies have explored whether they can process dialectal variation during the initial stages of L2 learning and from classroom exposure.

Studies on voice identification have reported a native-language bias, where listeners are more accurate at discriminating voices that speak their L1 as opposed to voices speaking an unfamiliar language (Goggin et al., 1991; Köster et al., 1995). Winters, Levi, and Pisoni (2008) found support for this L1 bias, while Wester (2012) observed that language familiarity is not relevant when listeners are asked to determine if two sentences were spoken by the same voice.

Studies with L2-learners indicate varying degrees of identification accuracy, but previous observations may have had several confounding factors. Köster et al. (1995) examined voice recognition in L2-learner groups with varying degrees of German proficiency. Participants had been in Germany for several months at the time of testing and were more accurate at identifying a German voice in a German-voice lineup than NS-English listeners without German knowledge. German NSs' and learner performance did not differ. Concluding that L2 proficiency does not affect voice identification, the authors revised their observations when a follow-up with L1-Spanish and L1-Chinese learners of German revealed both groups were significantly less accurate than the learners and NSs from the previous study (Köster & Schiller, 1997). The Spanish and Chinese groups had studied exclusively in Spain and China, without any immersion experience, and received German input in a classroom setting. The difference of immersion vs. classroom learning may thus have been a confound in those studies. Sullivan & Schlichting (2000) investigated L1-English learners of Swedish specifically in a classroom setting. First-year learners were more accurate at identifying a familiar voice in a lineup than a control group with no knowledge of Swedish, but accuracy among advanced learners was not higher than 1st-year

learners'. Advanced learners who had been abroad were also no more accurate than the lower-level learners, suggesting neither immersion nor higher proficiency are a relevant factor for improvement on voice identification. Here, task design may have also affected learner performance; the voice lineup was created using a voice imitator, deliberately increasing confusability.

THE PRESENT STUDY

In the above studies, L2 learners were generally less efficient at processing indexical cues. While some observed a learning curve for L2 learners, others found no indication of improvement. It is unclear how learner behavior in those studies was related to task effects (Sullivan & Schlichting, 2000), L2 immersion (Köster et al., 1995; Köster & Schiller, 1997) or L2 proficiency (Sullivan & Schlichting, 2000; Eisenstein, 1982). Those aspects are often underreported, limiting the conclusions to be drawn from previous investigations. If indexical information is found to be salient, learners might benefit from increased language variety and speaker variation in classroom material as previously suggested (Iverson, Hazan, & Bannister, 2005). If, however, learners learn to successfully interpret indexical information only after a certain amount of L2 exposure, classroom material should accommodate learners' capabilities.

The present study examined L2 learners' ability to distinguish dialect forms of German and German-speaking voices. In a cross-sectional investigation, L2 learners and German NSs completed a free classification task (FC). For learners, accuracy was predicted to increase with increasing L2 proficiency. Based on previous research, all groups were predicted to employ indexical cues to some degree (Clopper et al., 2006). However, for learners of German, less accurate processing skills were expected when compared to NS who have the L1 and familiarity advantage (Pisoni & Clopper, 2004; Winters et al., 2008). L2 learners also completed a language-background questionnaire, and a proficiency test (C-test) to allow for a more refined interpretation of the results. The results of this study contribute to our understanding of L2 learners' perception and processing of indexical information.

METHODOLOGY

Listeners

L1-American-English learners of German ($N = 27$) and German NS ($N = 5$) participated in this study. NS were recruited from regions where the investigated dialects were not spoken. L2 learners were recruited in German language classes at a Midwestern US university. Their parents were native speakers of English, but participants' residential history varied slightly. Eighteen had grown up in the Midwest, nine had moved to the Midwest region. All learners confirmed that their speech did not have salient dialect-specific characteristics. No participant reported a history of hearing or speech disorders, and all passed a hearing screening.

Talkers and Dialects

Talkers were five female German NS. Talker BD (age in years = 26) was from Southwest-Germany with Swabian as native dialect and had lived only in that region of Germany. The remaining four talkers had grown up and lived in the state of Saxony with Upper Saxon as their native dialect. Three (KN = 29, CX = 29, IK = 65) had never left the region for an extended period, while one talker (ES = 76) had permanently moved to the state of Brandenburg at age 25.

Stimulus Materials

Thirty different German words were chosen as stimuli. Talkers read a list of forty-six words once in their native dialect, and once in Standard German. Recordings were made with a Zoom H2n recorder in a quiet environment. All recordings were spliced in Praat and saved as individual sound files. For each speaker, three dialect (D) and three Standard German (Std) tokens were chosen. Swabian and Upper Saxon exhibit several phonetic characteristics that differ from Standard German (See appendix). Both merge aspirated and unaspirated German stops in syllable-initial context. Furthermore, word-final, unstressed syllables ending in <-r> are produced with a near-open, unrounded central vowel [ɐ] in Standard, but pharyngealized [oˤ] in both dialects (Ud Dowla Khan & Weise, 2013; Russ, 1990). Standard-German stimuli were judged by the author not to exhibit any salient features of variation and to conform to the phonetic forms described in the literature (Mangold, 2000).

Procedure

For the FC task, participants saw a Power Point slide with a 16x16 grid on a computer screen. Next to the grid were 30 consecutively-numbered boxes, connected to a sound file (one per stimulus). Listeners were instructed to listen to each file and drag the items onto the grid grouping them according to talker dialect and talker voice. To ensure participants focused on both target forms (voice and dialect), they were told that the same talker might speak different dialects and if they thought that was the case, they were encouraged to separate their groups further. Participants could listen to and rearrange the objects as often as they liked. Listeners then completed a background questionnaire in Qualtrics including a word-familiarity task, where they listened to the thirty stimuli from the FC task and gave the German spelling and English translation. Finally, a 30-minute, timed C-test was administered online. The task randomly assigned listeners five short paragraphs with blanks that had to be filled out.

ANALYSIS AND RESULTS

Based on the C-test, learners were assigned to the beginner (BEG, $N = 14$) or intermediate (INT, $N = 13$) group. The test automatically calculated learners' scores and assigned a proficiency level based on the *Common European Framework of Reference for Languages*. Out of 100, beginners (BEG) scored between 8-40 and Intermediate (INT) learners scored 43-80. Proficiency scores were strongly correlated with learners' course enrollment ($r = .97$, $p < .01$). For word familiarity, the Scheffé post hoc test revealed the beginners knew the task items significantly less often (Std: $F(3, 34) = 77.16$, $p < .01$, D: $F(3, 34) = 53.93$, $p < .01$) than the intermediates, which in turn were significantly less accurate than the NS (Std: $F(3, 34) = 77.20$, $p < .01$; D: $F(3, 34) = 53.93$, $p < .01$). On the background questionnaire only one BEG listener indicated having spent 2 months in Germany, while seven INT learners had been to Austria or Germany for an average of 3.7 months (range 1.5-12 months). No participant had been to a region where the dialects investigated here were spoken.

Listeners' Perceptual Space

ALSCAL. To investigate the perceived similarities and dissimilarities of the FC stimuli, a multidimensional scaling (MDS) analysis was conducted using SPSS 24's ALSCAL function. A 30x30 pairwise dissimilarity matrix was created for each group. Each cell indicated how many times listeners did not group an item together with another item. Based on the elbow in the stress

plot, a two-dimensional solution was deemed the best fit for all groups (BEG $r > .79$, stress = .27, INT $r > .83$, stress = .25, NS $r > .83$, stress = .25), suggesting that listeners perceived items as similar or dissimilar along two different dimensions (Figures 1, 2, 3).

Acoustic Analysis. To determine which perceptual dimensions were most salient for listeners, the fundamental frequency (f0) and the intensity (dB) of all stimuli were measured in Praat. Measurements were taken at 0%, 25%, 50%, and 75% of the stressed vowel in each stimulus. Dialect-specific dimensions, i.e. number of phonological processes in a D item as opposed to its StD form were tallied (0-3), and stimuli with frequently occurring processes were assigned a dichotomous value for absence or presence of the respective process (1=present, 0=absent). Speaker age was based on their biological age in years (1 = 25-35, 2 = 65, 3 = 76).

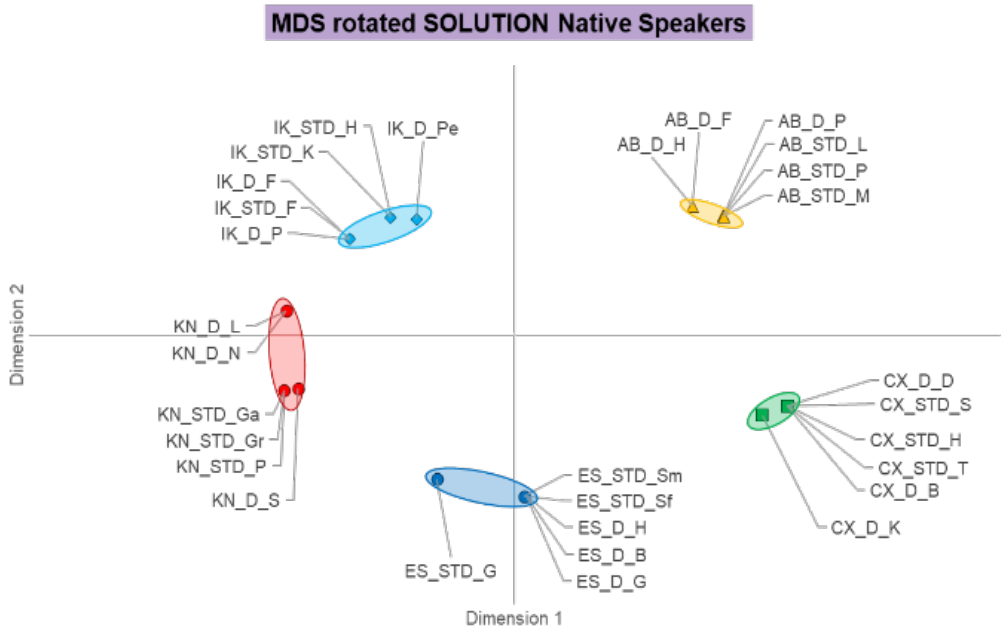


Figure 1. High variability training in the lab and in the language classroom.

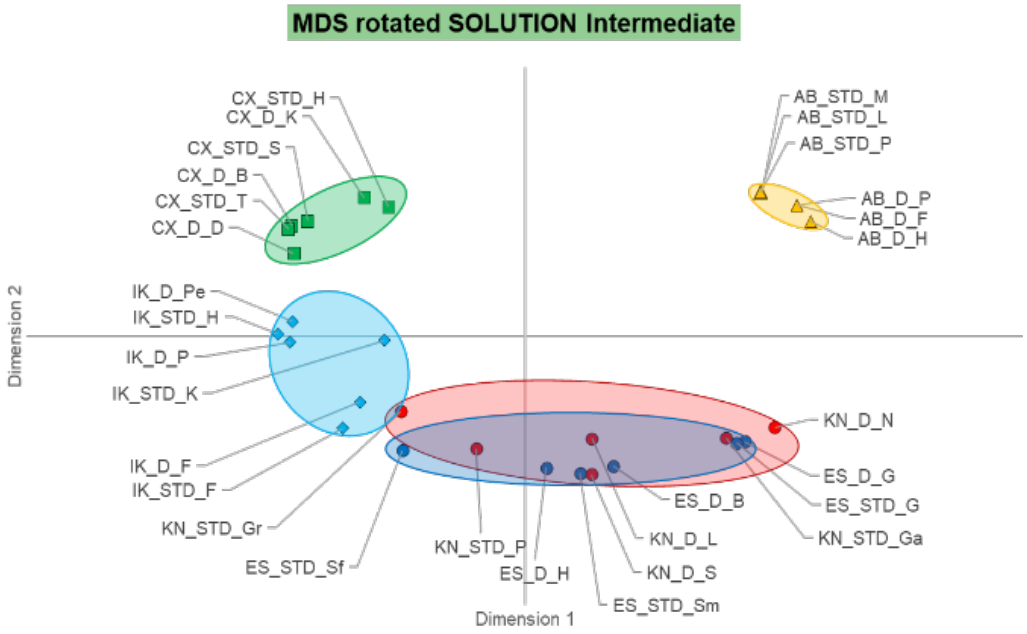


Figure 2. MDS rotated solution illustrating the perceptual space for intermediate level learners of German.

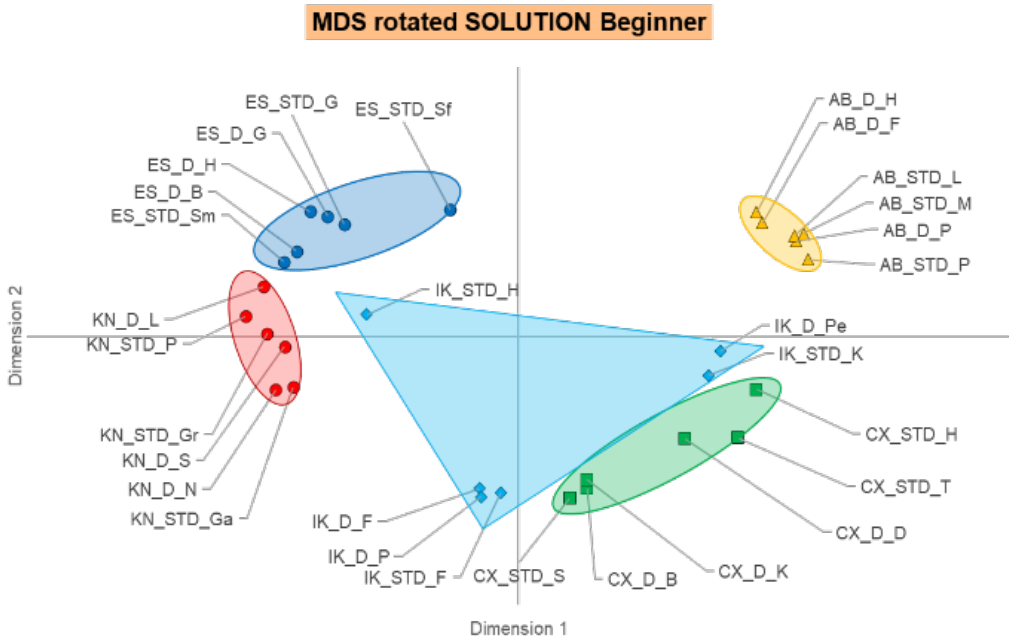


Figure 3. MDS rotated solution illustrating the perceptual space for beginning level learners of German.

Narrow groupings in the MDS solutions indicate that items were consistently grouped together, while wide-spread groupings indicate confusion about item group-membership. The nearly equidistant, narrow groupings in the NS-MDS solution reveal that listeners perceived items in other groups as very different. The wide-spread groupings in the INT- and BEG-MDS solutions show more overlap between stimuli, i.e. a larger number of stimuli was perceived as similar by learners where the NS made a clear distinction. After rotating the MDS solutions, a correlational analysis for acoustic or dialect-specific cues in Excel revealed the most salient dimension for all groups was *voice* (NS r = -.93, BEG r = .81, INT r = -.80) while neither group based their decisions on dialect-specific information (Table 1 below).

Table 1

Pearson correlation coefficient (R^2) for dialect-specific and acoustic cues calculated for all groups' rotated MDS solutions

Variable	**BEG**		**INT**		**NS**	
	Dim1	Dim2	Dim1	Dim2	Dim1	Dim2
dialect-specific						
variation (STD vs. D)	.07	-.01	.10	-.03	.0	-.04
acoustic measurements						
voice	**.81**	.27	-.27	**-.80**	**-.93**	.16
f0	**-.44**	.06	.05	.21	**.40**	.02
dB	-.27	**.55**	**.53**	.04	.35	**.44**
age	.19	**.50**	-.06	**-.61**	-.27	**.44**

For NS and BEG, *voice* was best interpreted as f0, while the INT group relied on speaker *age*. In the second dimension each group relied on either *age*, dB or a combination thereof. A series of linear regressions was carried out in SPSS to investigate if f0 or dB could predict talker or talker age. For f0, a significant effect was found for talker KN whose f0 values were significantly lower than the other talkers' (μ=165.6, r = .59, $F(4,25) = 3.4$, $p < .05$). Therefore, a certain f0 range corresponded strongly to KN tokens. Furthermore, dB was a predictor for the middle-aged talker (IK), the quietest (μ_{dB}=48.2, r = .78, $F(2,27) = 20.74$, $p < .01$) and for the oldest (ES), the loudest talker (μ_{dB}=60.5, r = .78, $F(2,27) = 20.74$, $p < .05$).

Listeners' Accuracy

The FC task also provides for objective accuracy measures, thus a difference score (DiffS) was calculated for each listener:

1) DiffS = (percent correct pairings) – (percent error pairings)

Percent correct and incorrect pairings are similar to 'hits' and 'false alarms' in signal detection theory (Clopper & Bradlow, 2009). Stimuli pairs were counted as correct pairing if both items

were spoken by the same talker and in the same variant (e.g. both D or both StD). Following Clopper & Bradlow (2009), the proportion of correct pairings was calculated out of the total possible number of correct talker pairings. Similarly, the proportion of mispairings was calculated out of the total possible number of incorrect pairings. A pair was counted as error if it had two different talkers, or the same talker but different variants grouped together. In Table 2, a low DiffS indicates that listeners frequently grouped items incorrectly, while a higher DiffS reflects more correct pairings. The maximal score on this task was 30.

Table 2

Between group comparisons for voice and dialect difference scores.

	BEG N = 14		INT N = 13		NS N = 5	
	M	SD	M	SD	M	SD
DiffS	-28.35**	20.05	-17.38	20.79	5	12.59

Note: *p < .05, **p < .01

An analysis of variance in SPSS revealed that the BEGs' DiffS was significantly lower than the NS groups' ($F(4, 34) = 3.59$, $p < .01$), while there was no significant difference in group means between the INT and NS groups (see Figure 4). In a correlational analysis, word familiarity was found to correlate with DiffS ($r = .73$), while LOR in the immersion setting did not ($r = .4$).

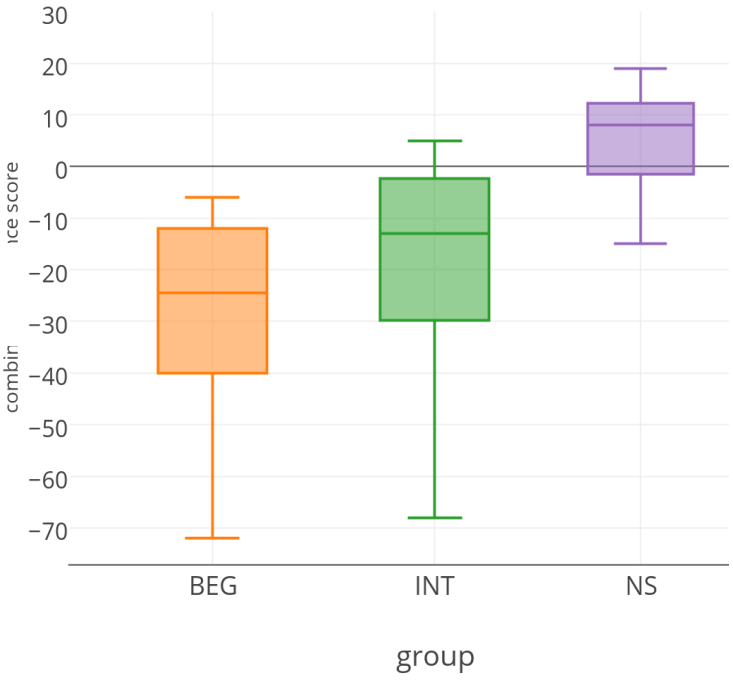


Figure 4. Boxplot for difference scores. This figure illustrates average DiffS for all groups with standard error bars.

DISCUSSION

This study investigated L2 learners' abilities to classify stimuli based on two indexical cues: talker voice and talker dialect. The classification results indicate that all groups could discriminate the five talkers. The NS performed more accurately than the learner groups, as expected from previous studies that reported an L1 advantage (Goggin et al., 1991; Winters et al., 2008). Classification accuracy, determined by difference score (DiffS), differed significantly for BEG and NS listeners. INT had a higher DiffS, but as much variation as BEG. Their MDS solution also revealed an overlap for talkers which was not observed in the NS group. Nevertheless, the INTs were not significantly different from NSs indicating a general trend for improvement with learner proficiency. However, this non-significance might be due to low statistical power from the small sample sizes. Larger samples from these populations are needed to draw more generalized conclusions.

The perceptual maps of the three groups revealed that listeners determined group membership based on f0 and dB, rather than on dialectal-specific information. All groups also relied on age. As the analysis revealed, age could be predicted based on dB level, where the middle-aged talker (IK) would have been perceived as the quietest speaker and the oldest talker (ES) as the loudest. f0 predicted which tokens were produced by KN. MDS solutions and DiffSs indicate that the groups used those cues with differing rates of success. BEG learners relied on f0 and reliably distinguished KN from the other speakers. They also used dB to classify tokens, which should have allowed them to identify IK from the other speakers. Yet, she was grouped with ES and CX and thus dispersed across the BEG perceptual map (Fig 3). The INT listeners relied only on dB, ignoring f0, which could explain KN and ES overlapping in their perceptual space (Fig 2).

The NS relied on the same cues as the learners yet were more successful at classifying stimuli. Previous studies that have investigated perceived correlates of age found that f0 and intensity are relevant, but other characteristics, e.g. laryngeal tension, air loss, and preciseness of articulation might be more important to distinguish voices from one another (Ryan & Burk, 1974). Thus, NS might have focused on a cluster of variables that the learners did not use. Since the analysis was limited to measurable acoustic cues, it may not have captured all voice-related features that listeners may have perceived as relevant. Thus, future investigations could benefit from a more precise voice assessment and post-task inquiries asking listeners to describe characteristics in the stimuli they deemed relevant for their decisions.

Despite more accurate groupings by the NS group, their DiffS did not approach ceiling levels and listeners only inconsistently grouped dialect vs. standard token. On the one hand, this might be because no NS had had significant exposure to the relevant dialects, a criterion found to increase dialect discrimination (Pisoni, 2004a). In that case, the NS results would support previous studies that found lower performance for unfamiliar language variation. On the other hand, the study design required listeners to focus on both voice and dialect for their decisions. The instructions for participants to focus on voice and dialect at the same time conflate several acoustic and perceived dimensions and complicate the interpretation of listeners' perceptual space. Future investigations should take care to separate them to obtain discrete perceptual spaces for voice and dialect classifications and allow for a more controlled investigation of acoustic dimensions that are salient to the NS and L2 learners.

While LOR in the L2 environment was not a good predictor for classification accuracy, word familiarity was shown to have influenced learners' results. BEG knew significantly fewer words

than INT who recognized significantly fewer words than NS. Stable lexical representations connected to a specific phonological form may have allowed NS listeners to compare the stimuli with the familiar form and make more consistent decisions.

Despite some limitations, this study sheds some light on classroom learner's abilities to process socio-indexical cues. The results indicate a learning curve for learners, illustrating that indexical processing might increase with proficiency. The interlanguage seems to also require stable representations for learners to reliably process indexical cues in the L2. Suggestions to increase the language input with highly variable stimuli, i.e. different voices, L2 variations, should take into consideration that learners' might initially not be equipped with the tools to process all types of information with equal precision and that exposure and practice could be necessary to allocate resources towards indexical information.

ABOUT THE AUTHOR

The author is a PhD candidate at Indiana University, Bloomington. Her research interest lies in second language acquisition and therein primarily in second language phonology. Previous and current works examine the perception and production of non-native segmental categories by second language learners and naïve listeners, as well as the establishing of phonological forms for lexical representations in second language learners. Her research also branches out into sociolinguistics and dialectology. Focusing on variation in German-speaking regions, she investigates the processing of dialectal variants by second language learners and second dialect acquirers, and she studies the potentials of incorporating indexical variation into the second-language classroom.

Correspondence concerning this article should be addressed to Franziska Kruger, Department of Germanic Studies, Global and International Studies Building 3103, 355 North Jordan Ave. Bloomington, IN 47405-1105, telephone: 812-855-1553, email: fkruger@indiana.edu

REFERENCES

Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech, 48*, 1-27.

Clopper, C. G., & Bradlow, A. R. (2008). Perception of dialect variation in noise: Intelligibility and classification. *Language and Speech, 51*, 175–198.

Clopper, C. G. & Bradlow, A. R. (2009). Free Classification of American English dialects by native and non-native listeners. *Journal of Phonetics, 37*, 436-451.

Clopper, C. G., Levi, S. V., & Pisoni, D. B. (2006). Perceptual similarity of regional dialects of American English. *JASA, 119,* 566-574.

Clopper, C. G. & Pisoni, D. B. (2004). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change, 16*, 31-48.

Cooke, M., García Lecumberri, M. L., Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *JASA, 123*, 414-427.

Eisenstein, M. (1982). A study of social variation in adult second language acquisition. *Language Learning, 32*, 367–391.

Eisenstein, M., & Verdi, G. (1985). The intelligibility of social dialects for working-class adult learners of English. *Language Learning, 35*, 287–298.

Goggin, J.P., Thompson, C.P., Strube, G., & Simental, L.R. (1991). The role of language familiarity in voice identification. *Memory & Cognition*, *19*, 448-458.

Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *JASA, 118*, 3267-3278.

Köster, O., & Schiller, N. O. (1997). Different influences of the native language of a listener on speaker recognition. *Forensic Linguistics, 4*, 18-28.

Köster, O., Schiller, N. O., & Künzel, H. J. (1995). The influence of native-language background on speaker recognition. In K. Elenius, & P. Branderud (Eds.), *Proceedings of the Thirteenth International Congress of Phonetic Sciences*, *4*, (pp. 306-309). Stockholm.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Mangold, M. (2000). Duden – Das Aussprachewörterbuch. Mannheim: Brockhaus.

Pisoni, D. B. (1997). Some thoughts on "Normalization" in speech perception. In K. Johnson, & J. H. Mullennix. *Talker variability in speech processing* (pp. 9-66). San Diego: Academic Press, Inc.

Russ, C. V. J., (1990). Swabian. In: C. V. J Russ (Ed.), *The dialects of Modern German. A linguistics survey* (pp. 337-363). Routledge: London.

Ryan, W. J., & Burk, K. W. (1974). Perceptual and acoustic correlates of aging in the speech of males. *Journal of Communication Disorder, 7*, 181-192.

Sullivan, K. P. H., & Schlichting, F. (2000). Speaker discrimination in a foreign language: first language environment, second language learners. *Forensic Linguistics, 7*(1), 95-111.

Ud Dowla Khan, S., & Weise, C. (2013). Upper Saxon (Chemnitz dialect). *Journal of the International Phonetic Association, 43*, 231–241, doi: 10.1017/S0025100313000145.

Wester, M. (2012). Talker discrimination across languages. *Speech Communication, 54*, 781-790.

Winters, S. J., Levi, S. V., & Pisoni, D. B. (2008). Identification and discrimination of bilingual talkers across languages. *JASA*, *123*, 4525-4538.

Appendix Table A1 - FC- and familiarity- task token in Standard and dialect

speaker	item	standard phonetic form	dialect phonetic form
BD	feather	[fe:tɐ]	[fæ:toˤ]
BD	home	[haɪmatʰ]	[hɔɪmətʰ]
BD	party	[pʰa:tʰi]	[pa:ti]
CX	cabin	[bu:də]	[bʉ:də]
CX	poet	[tiçtʰɐ]	[tiʃtoˤ]
CX	pillow	[kʰɪsn̩]	[kɪzn̩]
ES	blister	[pla:zə]	[pla:sə]
ES	belt	[gʏʁtʰl̩]	[gʏʁtl̩]
ES	hunger	[hʊŋɐ]	[hʊŋoˤ]
IK	fan	[fɛçɐ]	[fɛʃoˤ]
IK	whip	[pʰaɪ̯t͡ʃə]	[paɪ̯dʃə]
IK	powder	[pʰu:dɐ]	[pu:doˤ]
KN	lion	[lø:və]	[le:ʊə]
KN	needle	[na:dl̩]	[noˤ:dl̩]
KN	sleigh	[ʃlɪtn̩]	[ʃlɪdn̩]
BD	life	[le:bən]	
BD	coat	[mantl̩]	
BD	test	[pʰry:fʊŋ]	
CX	Satan	[za:tʰan]	
CX	cabin	[hʏtʰə]	
CX	trauma	[tʰraʊ̯ma]	
ES	grate	[kɪtʰɐ]	
ES	conductor	[ʃafnɐ]	
ES	mold	[ʃɪm̩l]	
IK	string	[fa:dn̩]	
IK	cave	[hø:lə]	
IK	cat	[kʰat͡sə]	
KN	beating	[pʰry:gl̩]	
KN	garden	[kaʁtn̩]	
KN	graphic	[kra:fikʰ]	

McCrocklin, S., Blanquera, K., & Loera, D. (2018). Student perceptions of university instructor accent in a linguistically diverse area. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 141-150). Ames, IA: Iowa State University.

STUDENT PERCEPTIONS OF UNIVERSITY INSTRUCTOR ACCENT IN A LINGUISTICALLY DIVERSE AREA

Shannon M. McCrocklin, Southern Illinois University, Carbondale
Kyle P. Blanquera, University of Texas, Rio Grande Valley
Deyna Loera, University of Texas, El Paso

As American universities promote globalization, they welcome Non-Native English Speakers (NNESs) as students and faculty. Yet, 40% of students are more likely to drop a class if taught by NNESs, and students have complained about NNES Teaching Assistants due to their accents. Importantly, exposure to non-native accents leads to greater tolerance of NNESs. This study seeks to understand how exposure to one non-native accent, in this case Spanish-accented English, impacts perception of unfamiliar accents. Preliminary findings show the majority (96%) of participants (n=107) had high levels of contact with native Spanish speakers. Participants were randomly presented one of three surveys using a verbal guise technique that featured three varieties of English (native, Spanish-accented, and Chinese-accented) as well as images of women representing three phenotypes. Students used Likert scales to rate the recorded instructor's personality and language ability. Across all ratings, the native English speaker recordings were rated most favorably. The phenotype presented played a relatively minor role. No noticeable difference was found between students' reactions to Chinese- or Spanish-accented speech. Finally, students showed a preference for taking a course with the Hispanic phenotype when paired with a native English accent.

INTRODUCTION

As American universities promote globalization, they welcome many staff and faculty who are Non-Native English Speakers (NNESs) (Jenkins, 2015). For many students, classes at universities are their first major exposure to people from other countries and to non-native speech. However, students do not always perceive this diversity as a positive aspect. Students have raised concerns regarding the qualifications of NNES Teaching Assistants due to their accents (Lippi-Green, 2012), and Rubin and Smith (1990) found that 40% of students are more likely to drop a class if taught by a NNES.

But Rubin (2012) pointed out that instead of focusing entirely on the NNESs, researchers should also examine the listeners; "any assessment of a speaker's speech performance could very well reflect nearly as much about the listener as about the speaker" (p. 11). In their study on perceptions of NNES Teaching Assistants, Rubin and Smith (1990) found that speaker proficiency was less determinant of listener comprehension than lecture topic or lecturer ethnicity. Further, when student listeners perceived high levels of accentedness, they judged those teachers to have poor teaching quality. Findings suggested that more work is needed to understand listener's judgments and exploration of ways to reducing linguistic stereotyping.

One particular experience that impacts the listener's perception of non-native accent is prior exposure to non-native accents. In a study on reverse linguistic stereotyping, Kang and Rubin

(2009) recruited 158 participants from the campus and local community to acquire a diversity of perspectives. Utilizing an Asian and Euro-American guise with a recorded academic lecture on galaxies, Kang and Rubin asked participants to fill in missing words from the lecture in a cloze test as a measure of comprehensibility. Participants also self-reported their own exposure to NNESs in a typical week and rated the teaching quality and accentedness of the recorded speakers. Kang and Rubin found that 18-23% of the variance in comprehensibility was tied to the perceived ethnicity of the speaker, but there were positive correlations between prior exposure to NNES accents (as well as the study of linguistics and languages) and greater tolerance of NNESs; students with prior exposure had more positive perceptions of the speakers and greater comprehension of the speech.

Therefore, while it may be important for universities to continue providing NNES instructor language and teaching training, it is also important to focus on the listeners. Undergraduates may benefit from training programs that promote listening skills and work to shift attitudes towards NNESs (Rubin & Smith, 1990). However, more information about the impacts of exposure to NNESs will be helpful in designing effective interventions or trainings for undergraduates. While exposure to non-native accents has been shown to reduce stereotyping and linguistic bias, it is not clear if exposure to speakers of a single language background will equally impact perceptions of all language backgrounds or if such exposure would have an outsized influence on only the perceptions of the specific language students have been exposed to.

Our research seeks to better understand language exposure by examining the following question: When studied among a population with increased exposure to one particular non-native accent, namely Spanish along the U.S. Mexico border, do participants display equally positive attitudes towards a familiar and unfamiliar non-native accent?

THE PRESENT STUDY

The study was conducted in the Rio Grande Valley (RGV) of South Texas. According to 2010 US Census Data, the majority of the population in the RGV is Hispanic (around 93%), around 75% of households report using Spanish as the first language, and about 25% of residents are foreign born (predominantly coming from Mexico). Further, research has shown that the RGV does not fit the more typical pattern of three-generation heritage language loss/shift to English; instead, bilingualism is commonly maintained into the fourth and fifth generations (Anderson-Mejías, 2005). Spanish is common in the RGV, and students are likely to be familiar with a Spanish accent in English speech. On the other hand, the Asian population of the RGV is less than 1%, and that number would notably include multiple national and linguistic backgrounds such as Chinese, Korean and Japanese (2011 US Census).

Survey

Data collection was managed through online surveys created in Qualtrics. To prepare the surveys, videos were created that matched an image of a woman with an audio clip of a female voice presenting course policies.

Videos

Images were collected of fourteen women that represented three different phenotypes: White/Caucasian, Hispanic, and Asian. Following procedures suggested by Kang and Rubin (2009), who discussed the impact of appearance on ratings, efforts were made to control for the

appearance of the women in the images. All of the women featured were in their late 20s to early 30s, wearing professional clothing and had dark hair (ranging from brown to black). The background of the images was a solid light grey color. The images were piloted for ratings of general attractiveness. Using Qualtrics to create a pilot survey, researchers included all images in a survey in which participants (n=56) from upper-level university courses and the community provided basic demographic data and rated images on a single feature, attractiveness, on a scale. Researchers used scores from the piloting to choose four images to include in the study (2 White/Euro-American, 1 Hispanic, and 1 Asian) that received the closest scores (were rated to have similar levels of attractiveness).

Recorded presentations of four sections of course policies were collected from five women that represented three different accents in English: native English (regionally standard), native Spanish (Mexican), and native Chinese (Mandarin). The non-native speech samples were piloted for similarities in level of accentedness primarily to match one of three Spanish speakers to the level of accentedness of the Chinese speaker. Similar to the piloting conducted for images, small samples of the audio recordings were included in a Qualtrics survey with a single scaled rating for accentedness. Only one Spanish speaker, a bilingual of Spanish and Macuiltianguis Zapotec, came close to the level of accentedness of the Chinese speaker. Although Spanish was reported to be her dominant language, her recordings were also sent to experts in the field of second language pronunciation to check whether her accent would represent a recognizable Spanish accent in English. Five of the six experts identified Spanish as a likely native language, allowing the recordings to be approved for use in the study.

The images and audio were then matched and used to create videos that showed the still image for the entirety of the audio recording. The White phenotype was matched in separate recordings to all three accents, while the Hispanic phenotype was only matched with the native-speaker and Spanish-accented speech and the Asian phenotype was only matched with the native speaker and Chinese-accented speech. This led to seven different video recordings.

Survey Questions

Three different surveys with four video matches on each, were created to prevent students from seeing reuse of images and recordings, which would alert students to variables being investigated This design also allowed for overlap across surveys and additional data points. The survey started with two questions regarding informed consent. Next, the four videos were presented. For each recording, five Likert scale items for rating intelligence, pleasantness, clarity/ease of understanding, fluency, and accentedness were included with a six-point range from strongly disagree to strongly agree. For each recording there was also one additional Likert scale item asking, “Do you think you would like to take a class with this instructor?” with a five-point range from definitely yes to definitely not. At the end of the survey, 11 questions were included to collect demographic information about the participants, including experience with non-native English speaking friends/family and instructors.

Procedure

Participants were recruited from freshman-level courses, including introductory composition courses and a course in university learning framework, at a large university in the RGV. To control the quality of sound for listening, students selected times to come into a monitored computer lab or office to take the survey. Participants were asked to halt consumption of any food or drink for the duration of the study and to dispose of any gum. Further, participants all

used the same type of headphones (Sony stereo, over-the-ear headphones- MDR ZX110) to complete the listening activity. Most participants needed 10-15 minutes to complete the activity.

Participants

The average age of participants (n=107) was 20.55 years. Participants were predominantly in their first year of college (58.9% reporting 1 or 2 semesters), while around 22% were in their second year, and around 20% were in the third (or later) year of college. Although students were recruited from introductory classes, the university has a large number of non-traditional students that may have transferred from other colleges or are completing degrees slowly due to family or work obligations. Of the participants, 66 were Female and 40 were Male, while one declined to report. The majority of participants had spent a substantial portion of their lives in the area; 64% reported having lived in the RGV for 16 years or more, while 26% reported 6-15 years, and 10% reported 1-5 years. The majority of participants reported being bilingual in English and Spanish. Figure 1 shows participants' self-perception of their language learning experiences.

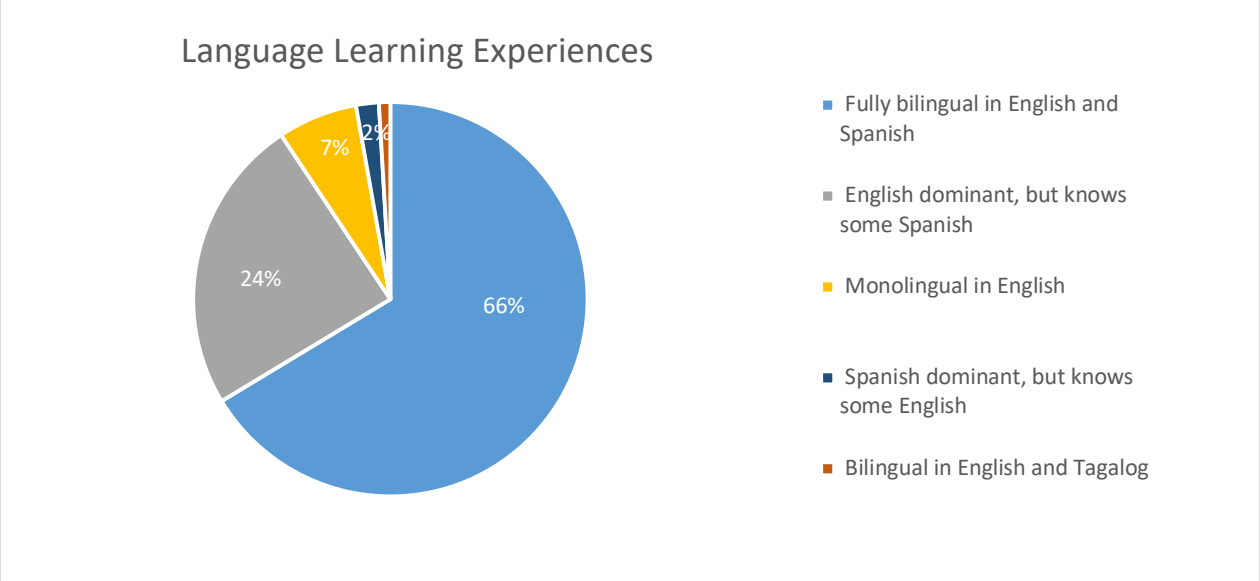


Figure 1. Participants' Primary Language Learning Experiences

About a quarter of participants (26.2%) also reported having studied another language, including Korean, French, American Sign Language, German, and Vietnamese.

Participants reported high levels of contact with Spanish first-language speakers. Almost all (96.3%) of the participants reported having close friends or relatives that spoke Spanish as their first language, while only 24% reported having close friends or family that spoke a language other than English or Spanish as their native language. For these 24% of participants, languages reported included French, Korean, Filipino, Akan, Chinese, Swedish, German, Arabic, and Vietnamese. Participants were also asked about their experiences taking classes in English that were taught by a non-native speaker in either high school or college. Figure 2 shows the participants' reported encounters with instructors of particular first languages, including "yes,"

"no," and "I'm not sure" if they had had an instructor with the particular language background. Most participants (76%) had taken courses in English with a native speaker of Spanish and were more likely to have had the instructor in high school than in college. Somewhat surprisingly, 45% of participants also reported having taken classes with instructors who spoke a language other than English or Spanish as their first language. Participants were more likely to have encountered the instructor in college than in high school. The languages of instructors included Filipino, Swahili, Chinese, German, Arabic, Russian, French, Korean, and Hindi. Taking friends,

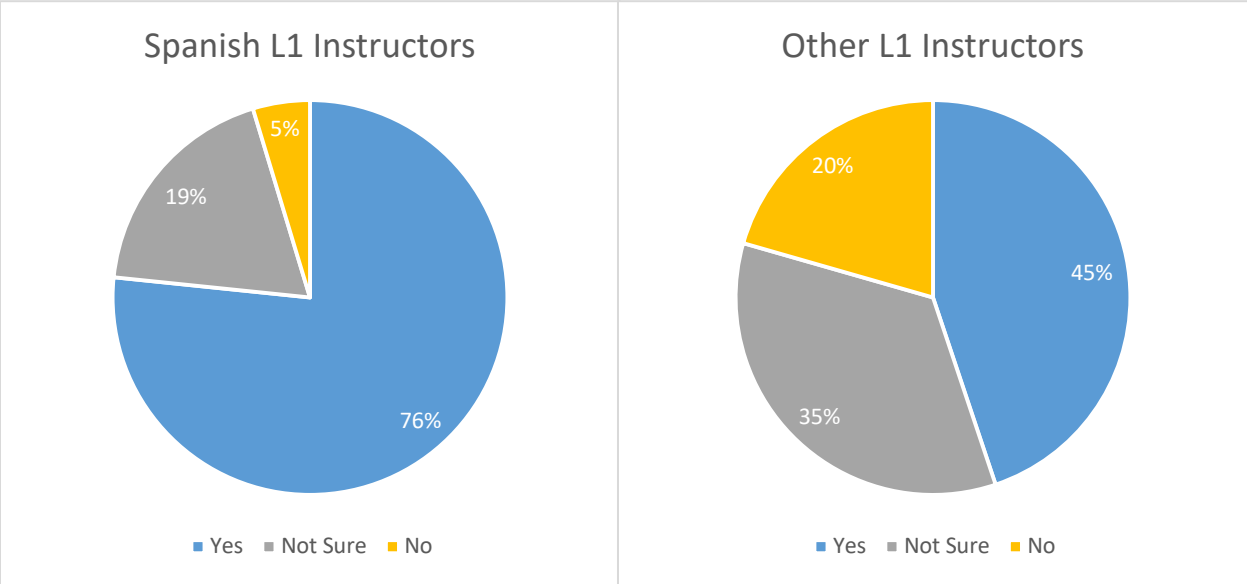


Figure 2. Participants' Exposure to Spanish and Other L1 Instructors

family, and instructors into account, only 6 participants (5.6%) had experience with native speakers of Chinese.

Analysis

The distribution of participants across the three surveys was reasonably even: 39 participants took survey 1, 33 participants took survey 2, and 35 participants took survey 3. Each recording/picture match existed on two surveys, which allowed for between 68-74 ratings for each video. An ANOVA was used to test for statistical significance in ratings between the various videos.

RESULTS AND DISCUSSION

The average ratings for the Likert scales following each video were calculated. On the first five items with a range of 0-5, higher scores indicated stronger agreement of the presence of targeted personality and language features. Averages on these items ranged from 1.70 disagree to 4.66. For the final Likert item with a range of 0-4, higher scores indicated a higher degree of interest and willingness to take the instructor represented by the video for a class. Averages on this item ranged from 1.51-3.05. Table 1 shows the average score for each video for each Likert scale item. The videos are labeled first by the phenotype represented (H- Hispanic, A-Asian, W-White/Caucasian) and then by the accent of the speaker (E- Native English, S-Spanish-accented

English, C- Chinese-accented English) so that, for example, WE represents the match of the White/Caucasian phenotype with the native English speaker accent.

Table 1

Average Score for each Likert Scale Item for each Video (picture/audio match)

Video/Match	Intelligent	Pleasant	Clear	Fluent	Accent	Would Take
WE	3.94	3.72	4.66	4.63	1.81	2.84
WS	3.69	3.03	2.13	2.44	4.13	1.49
WC	3.63	3.06	2.06	2.48	4.13	1.64
HE	4.09	3.85	4.54	4.42	1.86	3.05
HS	3.76	3.32	2.15	2.74	4.13	1.66
AE	4.15	3.63	4.64	4.59	1.70	2.99
AC	3.89	3.57	1.92	2.28	4.22	1.51

The ANOVA identified statistically significant differences between group means on all Likert scale items (p values ranging from .000-.048). A Tukey test post-hoc for each of the language feature Likert items illuminated significant differences between each of the native English accent pairings and each of the non-native accents (p values ranging from .000-.034). This difference existed for every native speaker-non-native speaker pairing, regardless of the phenotype represented. This result is not surprising as there are actual linguistic differences between the speech samples.

Conversely, there were no significant differences within a single accent speech sample (for example between HE and WE) or between the Spanish- and Chinese-accented speech. This indicates that students did not exhibit clear reverse linguistic stereotyping (judging the speech based on the image/phenotype presented), and students did not exhibit differences in judgments between the Spanish-accented speech that most (96.3%) were familiar with and the Chinese-accented speech that was unfamiliar to most (94.4%) of the participants.

For ratings of personality features (intelligence & pleasantness), the interactions were much more limited. There was a statistically significant difference between ratings of pleasantness for HE when compared to WC (p=.013) and WS (p=.004) and for WE when compared with WS (p=.013). Particularly notable in this finding is the lack of preference for the White/Caucasian image when paired with a non-native accent. Although the ANOVA identified a statistically

significant interaction for the item "intelligent" (p=.048), the Tukey test did not isolate any statistically significant pairings.

Although participants exhibited somewhat low levels of linguistic stereotyping in the measures of intelligent and pleasant, the Tukey test post-hoc on the measure of whether students would want to take a class with the instructor showed significant differences (all p values= .000) between each of the native English accent pairings and each of the non-native accents (for example, there was a statistically significant difference between the HE and HS pairings). There was, however, no difference between the phenotypes presented. This shows that despite exposure to Spanish-accented speech, students discriminated against L1 Spanish-speaking instructors, along with L1 Chinese-speaking instructors, while preferring the native English-speaking instructors.

Trends

Some trends in the data were also interesting although they were not statistically significant. In Table 2, each column shows the Likert item with each match (video) reordered by the average score with highest scores on top. To more clearly show patterns, the scores for accent have been flipped so that the lowest levels of accent are at the top. Further, the cell of each match has been recolored based on the accent represented (Blue=Native English, Grey= Spanish Accent, Dark Red= Chinese Accent)

Table 2

Videos Ranked by Highest Score per Likert Scale Item- Color Coded by Accent

Intelligent	Pleasant	Clear	Fluent	Accent	Would Take
AE	HE	WE	WE	AE	HE
HE	WE	AE	AE	WE	AE
WE	AE	HE	HE	HE	WE
AC	AC	HS	HS	WS	HS
HS	HS	WS	WC	WC	WC
WS	WC	WC	WS	HS	AC
WC	WS	AC	AC	AC	WS

From Table 2, it is possible to see that the scores for the native English accent were the highest across all items. Table 2 also highlights the fact that AC was the highest rated non-native speech when rated for intelligence and pleasantness. Participants may be drawing from cultural stereotypes that Asians possess superior intelligence and are inoffensive (if not pleasant). AC, however, received the lowest average scores when rated for clear, fluent, and accented. Participants also may have been drawing from expectations that the Chinese accent is more difficult to understand. Further, for the Chinese accent, students noticed language issues more when paired with an Asian face. However, this pattern did not exist for the Spanish-accent, and HS only averaged a lower score than WS on a single rating, accentedness.

Another interesting trend is that the White phenotype was rated lowest for intelligence when sorted by accent. AE and HE had higher averages than WE, while AC and HS had higher averages than WC and WS. This may be related to reverse linguistic stereotyping. When seeing the Asian or Hispanic phenotype, the participant may have expected an accent and, when none was present, may have given credit to intelligence for overcoming the accent. This may have also led to lower scores for the White phenotype when paired with an accent, because the White/Caucasian person is likely expected to have a native English accent.

Finally, participants were most likely to want to take a course with the Hispanic phenotype that was paired with a native English accent. This may relate to students feeling that they are more likely to relate to a person from a similar background, as 93% of residents in the RGV are Hispanic. The Hispanic phenotype was also preferred on the measure of would take within the non-native accents, ranking higher than WC, AC, and WS. The Hispanic phenotype also ranked quite high on the item pleasant for both the native and non-native accents.

Limitations

While this study provided insights into the reactions of listeners with prior exposure to non-native accents, there are limitations that need to be addressed. First, previous research has shown that stereotyping (linguistic and reverse linguistic) can lead to changes in intelligibility, or the actual ability of the listener to catch the message (Kang & Rubin, 2009). This was outside the scope of the current study, but would be useful to revisit with familiar and unfamiliar accents. Further, while a goal of the study was to attract mostly freshmen before they were exposed to numerous additional non-native accents while in college, the study included students in their second and third years of college. Around 45% of the participants had exposure to at least two non-native accents when entering the study. However, Chinese-accented speech was still unfamiliar to 94% of the participants. Finally, steps were carefully taken to ensure quiet and equivalent listening experiences for all participants across all three surveys. Unfortunately, noise issues in the labs (which only had flimsy barriers to prevent spread of noise) may have negatively impacted participants' abilities to carefully listen. This impacted all surveys equally, however, and the problem was solved early in the collection of data by moving the survey task into a quiet, private office.

CONCLUSION

While it is impossible to make exact comparisons with Rubin and Smith (1990) and Kang and Rubin (2009), the participants in this study who had frequent encounters with non-native English speakers (primarily Spanish L1) did not exhibit much reverse linguistic stereotyping, in which

the same audio file was rated significantly differently based on the image attached. It is notable that on the measure of pleasantness there was a dis-preference for the White phenotype when paired with a non-native accent. However, there were also some trends in the data, not statistically significant, that suggest that participants may have still been pulling from some cultural or linguistic stereotypes; the impact may have been minimized due to exposure to non-native accents.

The primary statistically significant differences were between the language feature ratings (clear, fluent, and accented) in comparisons of the native speaker and the non-native speaker audio samples. This is reasonable given that there were actual linguistic differences in the speech samples. However, it is concerning that despite prior exposure to non-native speech, participants still showed linguistic stereotyping and discrimination when it came to their judgments of whether they were interested in or willing to take a class with the instructor.

Finally, for none of the Likert scale items were there statistically significant differences between the Spanish- or the Chinese-accented speech. It seems that, for this group of participants, exposure to one non-native accent may have similar positive impacts on unfamiliar non-native accent tolerance. This makes training for listeners easier to accomplish as trainers do not have to work to expose students to every accent they may be likely to hear in order to increase tolerance. However, the findings suggest that exposure alone may not be enough to make students more willing to take courses with NNES instructors.

ABOUT THE AUTHORS

Shannon M. McCrocklin is an Assistant Professor at Southern Illinois University Carbondale in the Department of Linguistics. She teaches courses in TESOL, focusing on teaching speaking/listening skills in ESL as well as computer-assisted language learning. Her research focuses on second language acquisition of phonology and pronunciation teaching methods, including computer-assisted pronunciation training.

1000 Faner Dr., Room 3228
Carbondale, IL 62901
Shannon.mccrocklin@siu.edu
618-453-3428

Kyle P. Blanquera received his Bachelor's in English with emphasis on applied linguistics from The University of Texas—Pan American, and his Master's in English as a Second Language from The University of Texas Rio Grande Valley. His research interests include accent perception, sociolinguistics, applied linguistics, and language acquisition.

1201 W. University Dr, ELABS 329
Edinburg, TX 78539
Kyle.blanquera@gmail.com
956-665-3441

Deyna Loera is a graduate student at the University of Texas at El Paso studying for a Master's degree in Linguistics. She received her Bachelor's degree in English with a concentration in Linguistics and a minor in Sociology from the University of Texas Rio Grande Valley. Her fields

of interest include sociolinguistics, psycholinguistics, and language acquisition and is focused on research in those areas.

500 W. University, Liberal Arts Building 137
El Paso, TX 79968
Deyna.j.loera@gmail.com
915-747-5767

REFERENCES

Anderson-Mejías, P.L. (2005). Generation and Spanish language use in the Lower Rio Grande Valley of Texas. *Southwest Journal of Linguistics*, *24* (1), 1-12.

Jenkins, J. (2015). *Global Englishes: A resource book for students* (3rd Ed). New York, NY: Routledge.

Kang, O. & Rubin, D.L. (2009). Reverse Linguistic Stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology, 28*, 441-456. doi: 10.1177/0261927X99018001005

Lippi-Green, R. (2012). *English with an accent: Language, ideology and discrimination in the United States*. New York, NY: Routledge.

Rubin, D. (2012). The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgments and decisions. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2011. (pp. 11- 17). Ames, IA: Iowa State University.

Rubin, D. & Smith, K. (1990). Effects of accent, ethnicity, and lecture topic on undergraduates' perceptions of nonnative English-speaking teaching assistants. *International Journal of Intercultural Relations, 14* (3), 337-353. doi: 10.1016/0147-1767(90)90019-S

U.S. Census Bureau (2011). *Language Use in the United States: 2011*. Retrieved from https://www.census.gov/library/publications/2013/acs/acs-22.html.

Mora, J. C. & Levkina, M, (2018). Training vowel perception through map tasks: The role of linguistic and cognitive complexity. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 151-162). Ames, IA: Iowa State University.

TRAINING VOWEL PERCEPTION THROUGH MAP TASKS: THE ROLE OF LINGUISTIC AND COGNITIVE COMPLEXITY.

Joan C. Mora, University of Barcelona

Mayya Levkina, University of Barcelona

Recent research suggests that manipulating task design variables promotes attention to phonetic form during communicative interaction, leading to accuracy gains in production. This study investigated the effect of linguistic and cognitive complexity sequencing on the perception of a difficult L2 vowel contrast for L1-Spanish learners of English (/iː/-/ɪ/). The L2-English learners (*n*=81) were randomly assigned either to an experimental training group (EG, *n*=66) or a control group (CG, *n*=15). EG participants performed computerized map tasks that required them to give and follow directions to map locations that could only be successfully reached by accurately perceiving and producing the intended target contrast. Accuracy in perception was assessed before and after training through a categorical ABX discrimination task. The results revealed gains in discrimination accuracy and speed for EG participants, and treatment and experimental condition effects. Overall the results underscore the potential of task-based phonetic form-focused instruction for L2 pronunciation development.

INTRODUCTION

Most second language (L2) learners find pronunciation difficult to master and tend to exhibit strong L1 accents when speaking the L2, potentially causing comprehensibility problems when interacting with native speakers (Moyer, 1999). Similarly, L2 pronunciation remains a challenge for L2 teachers, who find pronunciation teaching difficult to integrate into their communication-oriented L2 classes (Darcy, Ewert, & Lidster, 2012). Still, empirical research on the effectiveness of task design and manipulation in L2 pronunciation teaching remains scarce (Gurzynski-Weiss, Long, & Solon, 2017), evidence of the difficulty in bridging the gap between research on L2 phonology and pronunciation teaching (Darcy, 2017). One of the difficulties of integrating pronunciation teaching into a communicative language teaching framework lies in designing pedagogic tasks that make a focus on phonetic form during student peer interaction essential for task completion (Mora & Levkina, 2017). Research has demonstrated that a focus on phonetic form and explicit corrective feedback may facilitate learners' awareness of L2 sound structure, possibly leading to benefits in L2 pronunciation development (Saito & Wu, 2014). One approach that is promising in bridging the gap between research and practice in L2 pronunciation teaching is task-based pronunciation teaching (TB*P*T), that is, the application of the tenets and principles of task-based language teaching (TBLT) to L2 pronunciation. A TB*P*T approach to pronunciation instruction generates pronunciation-focused communicative tasks by embedding a focus on phonetic form through task design manipulation (Mora & Levkina, 2017). Several recent studies have investigated the potential of TBLT principles and findings to extend beyond grammar and lexis to pronunciation.

In general the findings of these studies suggest that it is possible to manipulate the design parameters of communicative tasks to encourage a focus on phonetic form during communicative interaction, and that this leads to L2 pronunciation development, both at the level of segmental accuracy (Solon, Long, & Gurzynski-Weiss, 2017) and at the

suprasegmental level (McKinnon, 2017; Parlak & Ziegler, 2017). For example, Solon et al. (2017) tested the following TBLT predictions about increasing task complexity for L2 pronunciation. First, more complex tasks, as opposed to simpler tasks, should lead to a higher occurrence of pronunciation-based language related episodes (LREs) where students may engage in discussion about the target phonetic forms. Second, according to the *Cognition Hypothesis* (Robinson, 2003) more complex tasks should lead to larger gains in pronunciation accuracy than simpler tasks. Solon et al. set up a map task to be conducted in pairs (17 dyads; *n*=34) by L1-English intermediate level learners of Spanish containing minimal-pair street names targeting the 5 vowel phonemes of Spanish (/i/, /e/, /a/, /o/, /u/). Learners performed two map tasks varying in cognitive complexity, a simple one with few elements and a complex one containing many more elements. Although they found more (though not significantly more) pronunciation-focused LREs in the simple than in the complex task, contrary to their expectations, they found significant improvement in how accurately learners produced Spanish /e/ during the complex task, just as they had predicted.

As previous research within the Cognition Hypothesis (Robinson, 2001) had found for grammar and lexis, Solon et al. (2017) suggests that manipulating task complexity may benefit pronunciation accuracy. However, in their study map tasks were used to induce a focus on phonetic form through the use of minimal-pair street names contrasting vowels that were perceptually distinct (/a, i, e, o, u/) and embedded in a variety of target lexical items. It therefore remains unclear whether task complexity manipulation would have the same effect on confusable contrasting L2 vowels, such as English /iː/-/ɪ/ for L1-Spanish learners, or confusable vowels embedded in target word forms of varying linguistic complexity. In the current study we did not assess inter-group differences in segmental accuracy gains resulting from manipulations of task complexity. Instead, we kept cognitive complexity constant across participant groups by sequencing three map tasks in order of increasing cognitive complexity (simple, + complex, ++ complex), following Robinson's (2010) SSARC model, and manipulated linguistic complexity.

The present study addressed the following three research questions:

RQ1: Will a map task treatment be effective in improving L2 learners' perception of a confusable L2 vowel contrast (/iː/-/ɪ/)?

RQ2: Will linguistic complexity affect the effectiveness of the treatment?

RQ3: Will learners be able to generalize gains to new items and speakers?

THE PRESENT STUDY

According to Robinson's SSARC model, the main criterion in sequencing tasks is the manipulation of cognitive complexity along resource-dispersing and resource-directing variables (Robinson, 2010; Robinson & Gilabert, 2007). In our study, all participants, L1-Spanish learners of English, performed all three map tasks in the same order of cognitive complexity, but they were randomly assigned to one of three groups differing in the type of linguistic complexity of the task. We operationalized lingusitic complexity as the number of syllables a nonword contained. Three-syllable nonwords (*fadeetick-fadittick*) were defined as linguistically more complex than one-syllable nonwords (*deet-ditt*) because they contained more segmental material to process before the target stressed syllable. Perceiving the target vowel contrast /iː/-/ɪ/ in a three-syllable nonword was therefore expected to pose greater difficulty and to require greater attentional effort to learners than perceiving the same contrast in a corresponding 1-syllable nonword, which might lead to increased benefits in developing sensitivity to the contrast. Our aim, therefore, was twofold. First, we assessed the overall

effect of a sequence of map tasks on L2 segmental perceptual accuracy. Secondly, we assessed inter-group differences in perceptual accuracy gains as a function of the level of the complexity of the stimuli as defined above. Participants were assigned to one of three groups as a function of the complexity of the nonword items where the target English vowel contrast had been embedded: 1 syllable, 3 syllables, or a mixture of both. We chose the target English vowel contrast /iː/-/ɪ/ (e.g. *beat-bit*) because it has been shown to be difficult for Spanish learners of English to acquire, even at advanced levels of proficiency. The English vowel phonemes /iː/ and /ɪ/ are both perceptually mapped onto a single L1-Spanish phonemic vowel category /i/ (Cebrian, Mora, & Aliaga-Garcia, 2011). It is therefore a case of single category assimilation according to the Perceptual Assimilation Model PAM-L2 (Best & Tyler, 2007), which would predict considerable difficulty in acquisition for this English vowel contrast by L1-Spanish learners.

METHODS

The present study followed a pre-test > treatment > post-test design. All the participants took a perception test (AXB discrimination) and two production tests (delayed nonword and sentence repetition tasks) of the target vowel contrast /iː/-/ɪ/ in trained and untrained items. They also performed a map task test where they were asked to give and follow directions on a map. The participants in the experimental group, unlike those in the control group, did the treatment, administered over two weeks, consisting of perception and production familiarization tasks and three map tasks performed in order of increasing simple-to-complex cognitive complexity. Due to space limitations only the perception data is presented in the present paper.

Participants

The participants were 81 intermediate-to-advanced L1-Spanish learners of English. They had studied English in a foreign language instructional setting in Spain all through primary and secondary education and at the time of testing they were first-year university students starting a degree in philology. Participants were randomly assigned to an experimental group (EG: n=66) or to a control group (CG: n=15) that did not do any of the treatment tasks. Three participants in the experimental group were excluded from analysis because they missed the post-test. EG participants were further randomly assigned to one of three groups as a function of the type of stimuli they would be exposed to during the training tasks: EG1 (*mixed*) received a balanced mixed exposure to 1- and 3-syllable nonwords, EG2 (*simple*) were only exposed to 1-syllable nonwords and EG3 (*complex*) were only exposed to 3-syllable nonwords.

Stimuli and instruments

Four native speakers of standard Southern British English (2 female) recorded the speech stimuli for all tests and tasks. One female and one male voice were used for the familiarization pre-task and for the map tasks and half of the items in the ABX discrimination test. The other two voices were used for the other half of the items to test for generalization effects. The stimuli, with the target vowels /iː/ and /ɪ/ in stressed position, were either simple one-syllable CVC nonwords (*peef* vs. *piff*), or complex three-syllable CVˈCVCVC nonwords (*lapeefan* vs. *lapiffan*) created by adding phonotactically legal initial CV- and final -VC syllables to monosyllabic CVC nonwords.

ABX discrimination

The testing instrument was an ABX discrimination task. The ABX test contained 16 test nonword triads in 4 orders (ABB, ABA, BAA, BAB), i.e. 64 test trials, and 8 control nonword triads in 4 orders (32 control trials) based on 4 vowel contrasts (/i:/-/æ/, /ɪ/-/æ/, /i:/-/u:/, /ɪ/-/u:/) posing no discrimination difficulty. The test, with a total of 96 trials, was constructed so that in both test and control conditions there was always an equal number of 1- and 3-syllable nonwords, an equal number of trained and untrained nonwords, and an equal number of nonwords spoken by a male and a female voice. In each ABX triad, A and B were spoken by the same voice (either male or female) and X in a different voice (e.g., female if A and B were spoken by a male voice). Participants were instructed to decide whether the last item in a triad (X) was the same as the first (A) or the second (B) item as fast and as accurately as they could. ABX trials were presented with an inter-stimulus interval of 500 ms and an inter-trial interval of 2000 ms upon participant's response or 2500 ms after the onset of the last item in the triad if no response was provided. We obtained individual accuracy and reaction time (RT) scores for each participant across the main trial type conditions: test vs. control; 1 vs. 3 syllables; trained vs. untrained. RTs from wrong responses were excluded from analysis and the remaining RT data were screened by subject (and by condition) for RTs above or below 2.5 standard deviations (SDs) from the subject's mean RT.

Familiarization pre-task

For the familiarization pre-task and the three map tasks L2 learners were randomly assigned to one of three groups as a function of the linguistic complexity of the stimuli they were exposed to. The treatment pre-task consisted of an identification task containing the 48 "trained" nonwords in the ABX task described above. EG1 received exposure to 24 1-syllable (18 test + 6 control) and 24 3-syllable (18 test + 6 control) nonwords (36 test trials), EG2 was exposed to 48 1-syllable nonwords (18 test + 6 control x 2 repetitions; 36 test trials) and EG3 was exposed to 48 3-syllable nonwords (18 test + 6 control x 2 repetitions; 36 test trials). Nonword trials were randomly presented over headphones to participants, who were asked to identify the nonword they had heard by clicking on one of the 24 response buttons labelled in normal orthography. Minimal pairs (e.g. *piff* vs. *peef*) appeared side by side and were alphabetically distributed on the screen. The purpose of this task was to familiarize learners with the auditory and orthographic version of the nonwords they would encounter in the map tasks.

Map tasks

The participants in each of the three experimental groups (EGs) defined by the linguistic complexity of the stimuli (EG1: *mixed*; EG2: *simple*; EG3: *complex*) performed three computerized collaborative map tasks that differed only in the linguistic complexity of the nonwords used as street names (EG1: monosyllabic and trisyllabic; EG2: monosyllabic only; EG3: tri-syllabic only).

The tasks were performed in pairs so that first student A would give directions to student B, and then they would exchange roles. Students were sitting in front of two computer screens, one used for giving directions (1) and one for following directions (2). Student A would give directions to student B (following a marked itinerary on his map on screen 1) to pick a parcel located beyond a crossroad that led to two streets with minimal-pair names (e.g. *lapiffan* / *lapeefan*), each having a grey parcel icon. For example, student A would say: *now please pick the parcel you will find in "lapiffan" street*. Student B would need to decide whether it was the parcel in *lapiffan* street or the one in *lapeefan* street by clicking on it. In order to give participants feedback on accuracy, when clicked, parcels would go green if correct or red if wrong. Student A could always monitor what student B was doing on screen 2, but student B could not see student A's marked itinerary because his/her screen 2 would be switched off.

This situation raised clarification questions and repetition requests involving the contrasting street names, which generated pronunciation-focused LREs. In the maps used in the map tasks, the names of the streets were those used in the familiarization pre-task and were clickable, so that both student A and B could click on them to hear them over headphones. The map task sessions were digitally recorded (44.1kHz, 16-bit) onto Marantz PMD660 recorders separately for each student via Shure SM58 microphones.

RESULTS

The first aim in the current study was to assess perceptual gains in L2 learners' sensitivity to the /i:/-/ɪ/ contrast resulting from the treatment (a sequence of map tasks). We first examined mean differences between test and control trials. At pre-test control trials in the ABX discrimination test obtained significantly higher accuracy scores (proportion of correct responses) than test items did (see Table 1; EGs: $t(6334)=-16.52$, $p<.001$; CG: $t(1438)=-8.97$, $p<.001$) and RTs were significantly faster on control trials than on test trials (EG: $t(4397)=5.97$, $p<.001$; CG: $t(1031)=3.93$, $p<.001$). The same pattern of results was obtained at post-test, suggesting that the test contrast /i:/-/ɪ/ was difficult to discriminate, with mean 0.64-0.70 proportion correct scores. We next examined perceptual gains for the test trials only for all participant groups. In general, results showed differences in the expected direction, with post-test scores reflecting significant improvement in response accuracy ($t(8372)=-5.97$, $p<.001$) and speed ($t(5376)=9.28$, $p<.001$) for the EGs. As expected, improvement in accuracy did not reach significance ($t(1918)=-1.41$, $p=.158$) for the CG,, but response speed did ($t(1261)=4.32$, $p<.001$).

Table 1

Accuracy and RT scores by item type and testing time.

Group	Item Type	Test	Accuracy			RT		
			N	*M*	*SD*	*N*	*M*	*SD*
EGs	Test	Pre	4224	.64	.48	2664	1156.08	363.94
		Post	4160	.70	.46	2858	1069.25	331.33
	Control	Pre	2112	.84	.37	1735	1089.09	362.96
		Post	2080	.90	.30	1831	1001.76	328.46
CG	Test	Pre	960	.66	.47	620	1160.70	332.17
		Post	960	.69	.46	643	1084.15	295.93
	Control	Pre	480	.88	.33	413	1079.50	315.36
		Post	480	.92	.28	431	1013.72	320.34

Secondly, we assessed experimental inter-group differences in perceptual accuracy gains as a function of the level of the complexity of the stimuli for test trials only. Gains in accuracy and speed were larger for experimental subjects in the simple (EG2) and complex (EG3) conditions than for subjects in the mixed condition (EG1), suggesting that exposure to items of varying complexity might have been detrimental to accuracy gains (Figure 1). In order to test this we submitted the accuracy and speed scores to two-way ANOVAs with *Test* (pre-test, post-test) as a within-subjects factor and *Group* (mixed, simple, complex) as the between-subjects factor. The results showed, for accuracy, a significant main effect of *Test*

($F(1,60)=37.99$, $p<.001$, $\eta^2=.388$), but neither the main effect of *Group* ($F(2, 60)=.141$, $p=.869$, $\eta^2=.005$) nor the *Test* x *Group* interaction ($F(2, 60)=2.47$, $p=.092$, $\eta^2=.076$) reached significance. For speed, the same pattern of results was found, with a significant main effect of *Test* ($F(1,60)=28.86$, $p<.001$, $\eta^2=.325$) and a non-significant main effect of *Group* ($F(2, 60)=2.23$, $p=.115$, $\eta^2=.069$). The *Test* x *Group* interaction was not significant ($F(2, 60)=.391$, $p=.678$, $\eta^2=.013$). These results show that the treatment was effective in improving L2 learners' perception of the L2 vowel contrast /iː/-/ɪ/ (RQ1).

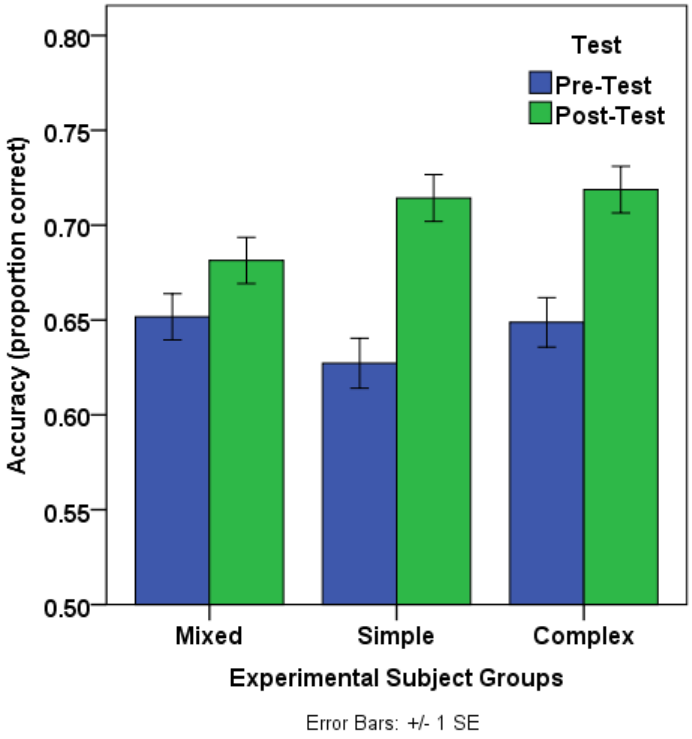


Figure 1. Proportion of correct responses by Experimental Group.

Gain sizes were larger for the groups exposed to the simple and complex treatment types (EG2 and EG3) than for the group exposed to a mixed treatment (EG1), but such group differences based on the linguistic complexity of stimuli did not reach significance (RQ2). This was further confirmed by a series of ANOVAs on gain scores (post-test minus pre-test) for accuracy and response speed (Figure 2). For accuracy gains the main effect of *Group* approached significance ($F(2,60)=2.479$, $p=.092$, $\eta^2=.076$), and was not significant for response speed ($F(2,60)=.391$, $p=.678$, $\eta^2=.013$). Accuracy benefits therefore patterned as *Simple* > *Complex* > *Mixed*, whereas RT benefits patterned as *Simple* = *Complex* > *Mixed*, but non-significantly.

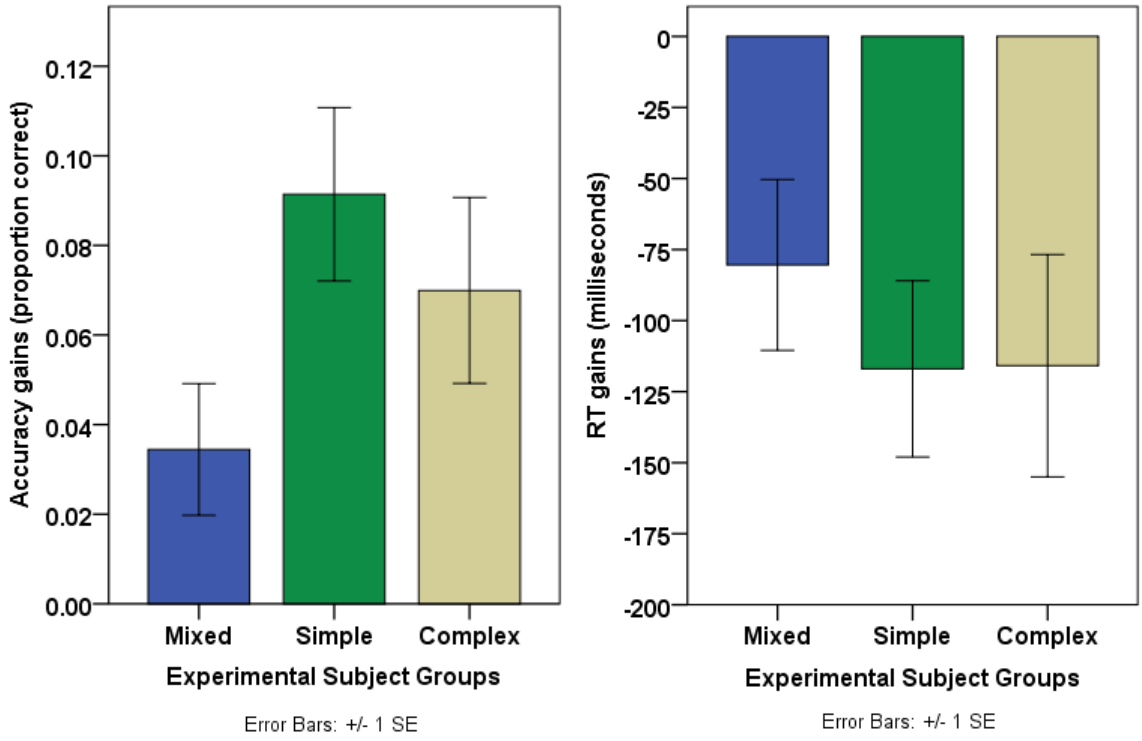


Figure 2. Accuracy and speed gains by experimental group.

Because half of the test items in the ABX discrimination test were monosyllabic (e.g. *piff*) and half were tri-syllabic (e.g. *lapiffan*), and participants had been assigned to different linguistic complexity treatment groups based on whether they were exposed to monosyllabic, tri-syllabic or both types of stimuli, it is likely for the three experimental groups to differ from one another in how accurately and how fast they could correctly identify 1-syllable and 3-syllable test items. In order to explore this possibility we analysed participants' performance separately by test item type (1 or 3 syllables). As Figure 3 shows, EG1 (mixed) obtained the smallest gains in both 1- and 3-syllable nonwords, whereas EG2 (simple) obtained the largest gains on 1-syllable nonwords, which appeared to be as large as those they obtained on 3-syllable nonwords, whereas EG3 (complex) obtained larger gains on 3-syllable words than on 1-syllable words.

We tested these differential gains through two ANOVAs with subject *Group* (Mixed, Simple, Complex) as the between subjects factor and *Test* (pre-test, post-test) and *Trial Type* (1-syllable trials, 3-syllable trials) as within subjects factors, one for response accuracy (proportion correct), and one for response speed (RTs in milliseconds). For both accuracy and speed the main effects of *Group* were not significant ($F(2,60)=.141$, $p=.869$, $\eta^2=.005$; $F(2,60)=2.239$, $p=.115$, $\eta^2=.069$; respectively) whereas the main effects of *Test* ($F(1,60)=37.99$, $p<.001$, $\eta^2=.388$; $F(1,60)=28.86$, $p<.001$, $\eta^2=.325$) and *Trial Type* ($F(1,60)=15.91$, $p<.001$, $\eta^2=.210$; $F(1,60)=246.83$, $p<.001$, $\eta^2=.804$) were significant. None of the interactions reached significance. This suggests that all groups significantly improved from pre- to post-test on both 1- and 3-syllable trials, even if participants generally found 3-syllable trials were harder and slower to discriminate[1]. However, within-group *t*-tests

[1]RTs were measured from the onset of the last nonword in an ABX triad for both 1- and 3-syllable trials. On average RTs on 3-syllable trials were 128 ms (pre-test) and 105 ms (post-test) slower than 1-syllable trials, whereas the initial unstressed syllable in 3-syllable nonwords was approximately 100 ms long. Therefore, response speed differences as a function of linguistic complexity cannot be attibuted to increrased difficulty or complexity, but to the extra response time needed to respond due to the initial syllable before the target contrast.

contrasting 1-syllable vs. 3-syllable mean accuracy and speed scores failed to reveal significance in any of the differences observed. In addition, the size of gains in accuracy ($t(62)=.066$, $p=.948$) and speed ($t(62)=.1.74$, $p=.0.81$) were comparable for 1-syllable and 3-syllable test-trials.

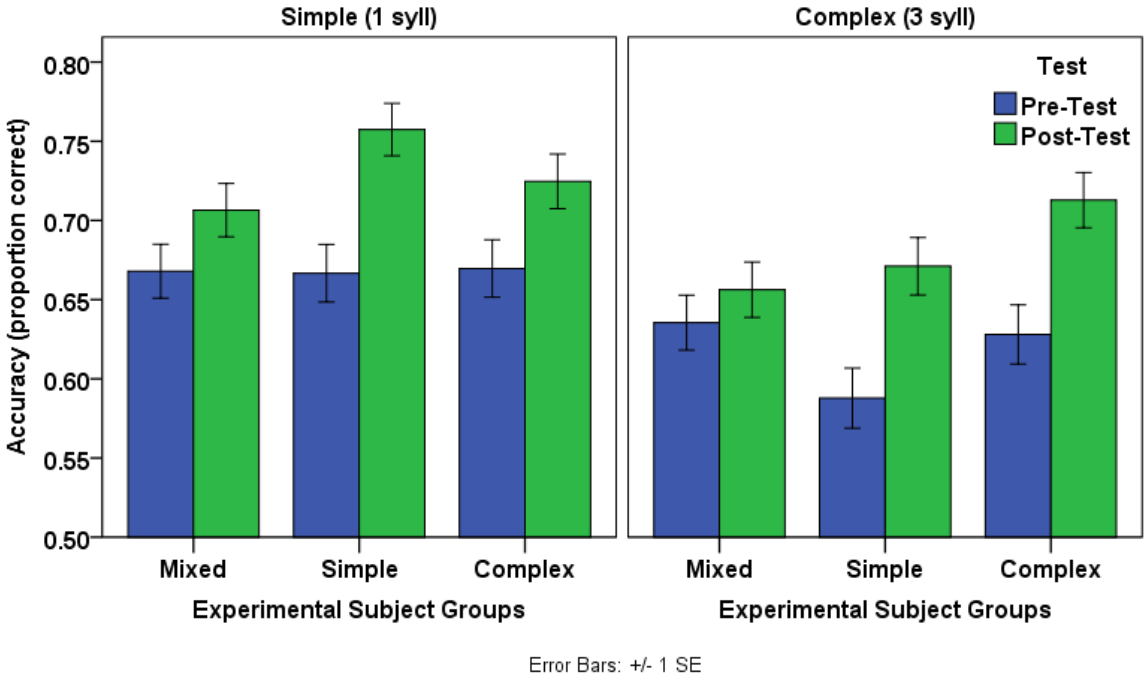


Figure 3. Pre-test and post-test accuracy scores by EG and Trial Type.

Finally, in order to test for generalization effects (RQ3), we assessed the extent to which the general significant gains in accuracy and speed found applied equally to trained and untrained test items. Test items containing untrained nonwords (new nonwords and nonwords spoken in by a new voice) significantly improved in accuracy ($M=.63>.69$; $t(4187)=-3.86$, $p<.001$) and speed ($M=1158>1067$; $t(2617)=6.69$, $p<.001$) from pre-test to post-test. The test items produced in a new voice were also found to improve significantly in accuracy ($M=.64>.69$; $t(4187)=-3.81$, $p<.001$) and speed ($M=1116>1028$; $t(2670)=6.57$, $p<.001$). These results show that participants were able to generalize gains to new contexts ("nonwords") and to new speakers ("voices") suggesting that they had learned to apply the discrimination skills acquired through the training to novel test items. As shown in Figure 4, gains for untrained items were only slightly smaller than those for trained items, as expected, and occurred in all experimental groups.

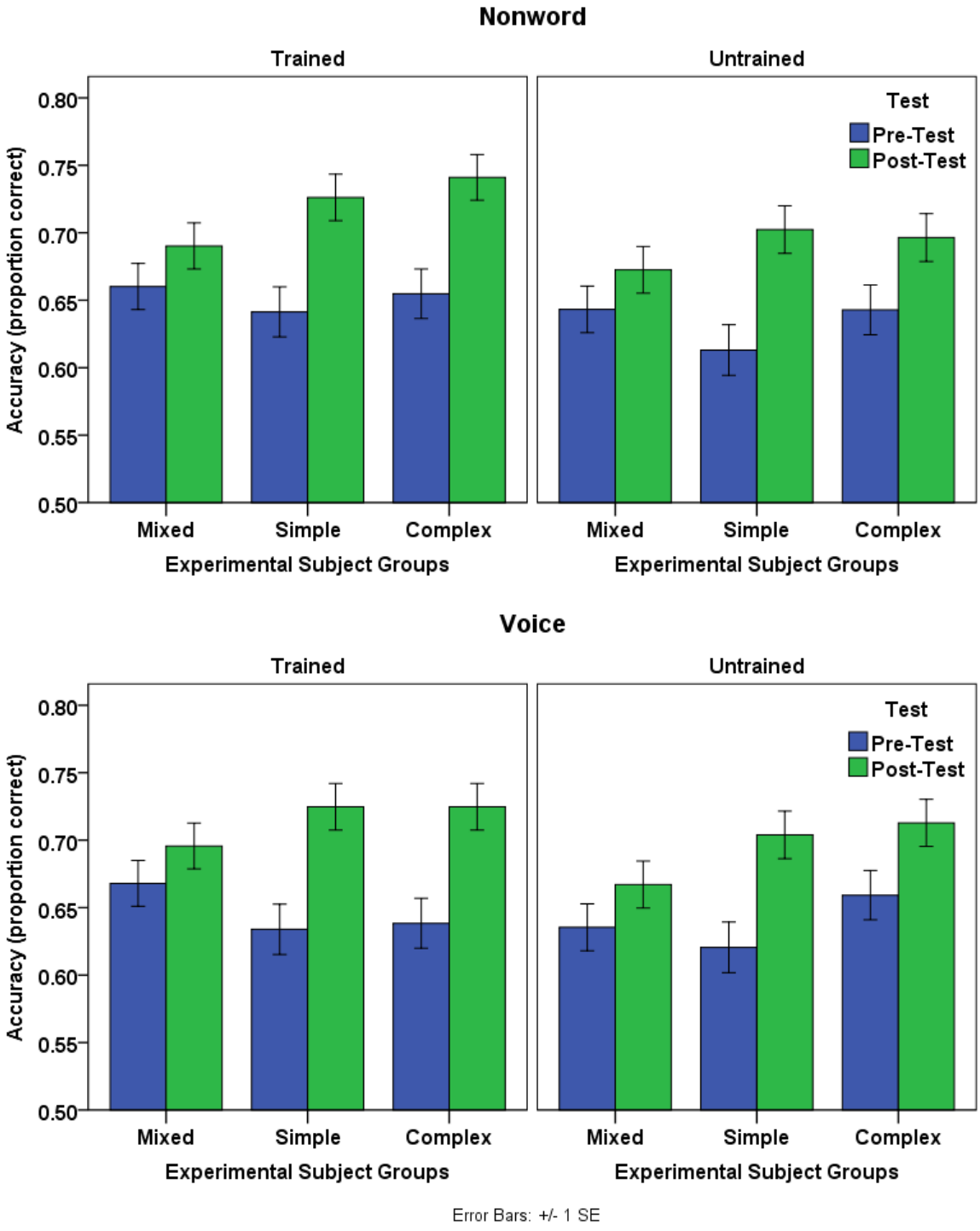


Figure 4. Pre-test and post-test accuracy scores by experimental group and item type.

DISCUSSION AND CONCLUSIONS

The present study has shown that it is possible to improve L2 learners' perception of a difficult vowel contrast through a series of collaborative map tasks performed in pairs. Although we did not manipulate cognitive complexity experimentally, the sequencing of the map tasks in terms of increasing cognitive complexity led to robust gains in discrimination accuracy and speed. Such gains appeared to be consistent across experimental groups receiving exposure to nonwords that differed in linguistic complexity (operationalized as

syllable number). In general, it was harder to discriminate the target /iː/-/ɪ/ contrast in tri-syllabic than in monosyllabic nonwords, but experimental groups receiving exposure to only one type of nonword (EG2 and EG3) significantly improved on both types, whereas the group receiving exposure to both types of nonwords appeared to improve the least (albeit not significantly so). We can therefore conclude that linguistic complexity did not significantly affect the effectiveness of the treatment, although the size of the accuracy gains obtained suggest that exposure to monosyllabic nonwords enabled learners to obtain slightly larger gains on all nonword types. The results also showed that learners were able to generalize gains to new items and speakers, indicating that improvement in discrimination accuracy was robust and suggestive of learning having taken place. Overall the findings of the present study suggest that manipulating task design variables to include a focus on phonetic form is an effective pedagogical strategy in generating pronunciation-focused LREs and in improving learners' accuracy in segmental perception while performing an interactive communicative task. It also suggests that exposing learners to simple stimuli in terms of phonological structure in pronunciation tasks may be more beneficial in learning phonological contrasts than exposing them to more complex stimuli. Further research should examine the effectiveness of manipulating task design in pronunciation tasks to include simple-to-complex task sequencing based on linguistic (phonological) complexity in combination with task sequencing based on cognitive complexity.

The current study presents a number of limitations that require further investigation in future research on task-based pronunciation teaching. One crucial issue to be investigated further is the extent to which the treatment effects observed were mainly due to the interactive map tasks or to the familiarization pre-tasks. This would involve recruiting a second control group that would do the pre-tasks, but not the interactive map tasks. It would also be important to include a delayed post-test to test for differential retention of gains in experimental and control groups.

ACKNOWLEDGMENTS

We would like to thank Ingrid Mora-Plaza, Diana Moreira de Oliveira and Natalia Wisniewska for their help in data collection.

ABOUT THE AUTHORS

Joan C. Mora is associate professor in the Department of Modern Languages and Literatures and English Studies in University of Barcelona (UB). His research has examined the role of contextual and individual factors in the development of L2 speech and oral fluency, and the acquisition of L2 phonology.

Contact information: mora@ub.edu, https://joancmora.weebly.com/

Mayya Levkina is assistant professor in the Department of Modern Languages and Literatures and English Studies in University of Barcelona (UB). Her research focuses on the development of different aspects of English as an L2 (e.g. pronunciation, pragmatics) through TBLT as mediated by working memory capacity, attention and other cognitive skills.

Contact information: mayya.levkina@ub.edu

REFERENCES

Best, C., & Tyler, M. (2007) Non-native and second language speech perception. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning* (pp. 15-34). Amsterdam: John Benjamins.

Cebrian, J. Mora, J. C., & Aliaga-Garcia, C. (2011) Assessing crosslinguistic similarity by means of rated discrimination and perceptual assimilation tasks. In Dziubalska-Kołaczyk, K., Wrembel, M., Kul, M. (ed.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010, Volume I* (pp.41-52). Frankfurt am Main: Peter Lang.

Darcy, I. (2017) A psycholinguist walks into a classroom…: A road-map for bridging research and practice in pronunciation teaching. Keynote presentation given at the *9th Annual Conference Pronunciation in Second Language Learning & Teaching* (PSLLT), September 1-2, 2017, University of Utah, Salt Lake City, Utah, USA.

Darcy, I., Ewert, D., & Lidster, R. (2012). Bringing pronunciation instruction back into the classroom: An ESL Teachers' pronunciation "toolbox". In J. Levis & K. LeVelle (Eds.), *Proceedings of the 3rd Pronunciation in Second Language Learning and Teaching Conference* (pp. 93-108). Ames, IA: Iowa State University.

Gurzynski-Weiss, L., Long, A. Y., & Solon, M. (Eds.) (2017). TBLT and L2 pronunciation: Do the benefits of tasks extend beyond grammar and lexis? [Special issue]. *Studies in Second Language Acquisition,* 39(2).

McKinnon, S. (2017). TBLT instructional effects on tonal alignment and pitch range in L2 Spanish imperatives versus declaratives. *Studies in Second Language Acquisition*, 39(2), 287-317.

Mora, J. & Levkina, M. (2017). Task-based pronunciation teaching and research: key issues and future directions. *Studies in Second Language Acquisition, 39*(2), 381-399.

Moyer, A. (1999). Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition*, *21*(1), 81-108.

Parlak, Ö., & Ziegler, N. (2017). The impact of recasts on the development of primary stress in a synchronous computer-mediated environment. *Studies in Second Language Acquisition*, 39(2), 257-285.

Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22,* 27-57.

Robinson, P. (2003). The Cognition Hypothesis, task design, and adult task-based language learning. *Second Language Studies*, 21, 45-105.

Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In M. Putz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind.* (pp. 243-268). Amsterdam: John Benjamins.

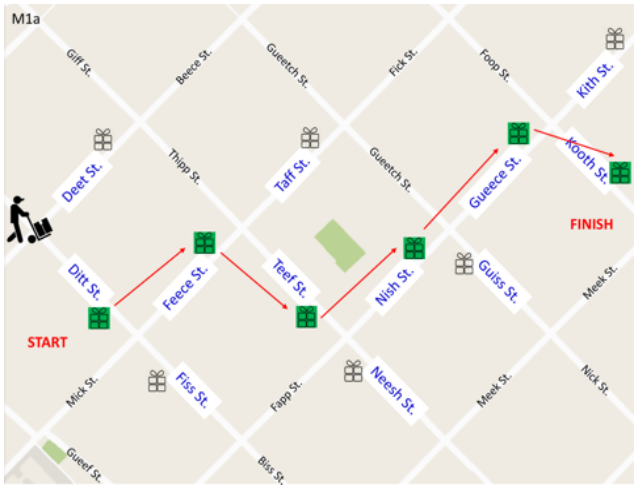
Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176.

Saito, K., & Wu, X. (2014). Communicative focus on form and L2 suprasegmental learning: Teaching Cantonese learners to perceive Mandarin tones. *Studies in Second Language Acquisition, 36*, 647-680.

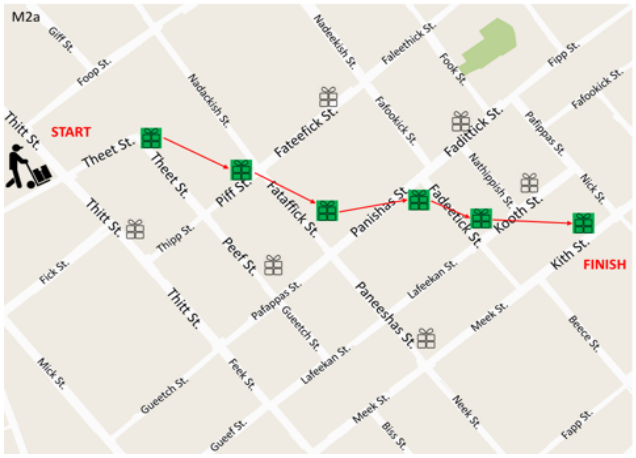
Solon, M., Long, A. Y., & Gurzynski-Weiss, L. (2017). Task complexity, language-related episodes, and production of L2 Spanish vowels. *Studies in Second Language Acquisition, 39*(2), 347-380.

APPENDIX: SAMPLE MAP TASKS

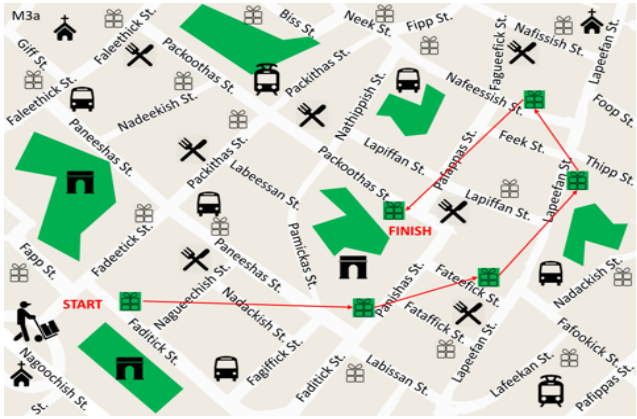
Simple Map Task



+Complex Map Task



++ Complex Map Task



Ruellot, V. (2018). Can Pepé le Pew help? Stereotypical accent and French pronunciation learning. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 163-173). Ames, IA: Iowa State University.

CAN PEPÉ LE PEW HELP? STEREOTYPICAL ACCENT AND FRENCH PRONUNCIATION LEARNING

Viviane Ruellot, Western Michigan University

This study examines the impact of stereotypical accent-based training on the acquisition of second language (L2) French pronunciation. Research suggests that L2 accent imitation in the native language (L1) benefits L2 pronunciation acquisition (e.g., Everitt, 2015; Rojczyk, Porzuczek & Bergier, 2013). This study seeks to contribute to the research in this emerging pedagogical approach and reports on its benefits for the pronunciation improvement of /ʁ/ in L2 French. Over three weeks, eleven American students received training in select characteristics of French accent and practiced their pronunciation by imitating models in three experimental groups: one in which the models spoke English with a stereotypical French accent (n=4), another where they spoke it with an authentic (i.e., non-stereotypical) French accent (n=4), and a third one modeled by native speakers of French speaking French (n=3). Students were recorded reading texts and describing pictures before and after practice. Findings from French native speaker ratings indicate no significant pronunciation improvement of French /ʁ/. Results are discussed in terms of length of training and number of features involved.

INTRODUCTION

After sitting on the 'back burner' of second language (L2) pedagogy for the first decade of the Communicative Approach, pronunciation instruction has resurfaced in language courses since the mid-1980s (Chun, 1991), accompanied by an ever-growing body of research validating its benefits (Lee, Jang & Plonsky, 2015; Saito, 2012). Although an expanding number of pronunciation instruction approaches involves technology (for a review, see Derwing & Munro, 2015 and Thomson & Derwing, 2015), many of them remain based on the traditional practice of listening to—and repeating after—models speaking the L2. These imitation tasks include shadowing, i.e., imitating a model simultaneously or after a slight delay, and mirroring, i.e., imitating exactly the speech and body movements of a model (Derwing & Munro, 2015). An emerging instructional approach has learners practice their pronunciation by speaking their native language with a second language accent. It is believed that the exclusive focus on pronunciation, unencumbered by the need to process other aspects of the language, may lead learners to develop awareness of cross-language segmental and suprasegmental differences and reproduce them in their L2 speech (Everitt, 2015). There is a growing body of experimental research investigating the impact of L2 accent imitation in the L1 on the production and perception of pronunciation features in L2 English (Everitt, 2015; Flege & Hammond, 1982; Mora, Rochdi & Kivistö-de Souza, 2014; Sypiańska & Olender, 2016) and in L2 French and German (Neuhauser, 2011). To our knowledge, Everitt (2015) is the only empirical research investigating the use of L2 accent imitation in the L1 as a pronunciation teaching and learning tool. This study—which is part of a larger one—builds on Everitt's, by investigating the benefits of this approach on the development of French /ʁ/.

Background

Imitation is at the core of human learning. By replicating the movements, behaviors and vocalizations of others, imitation contributes to the acquisition of many skills (Hauser, 1996; Nagell, Olguin & Tomasello, 1993; Zentall & Akins, 2001), including language. Infants acquire their native language by imitating speech sounds in their environment (Kuhl & Meltzoff, 1996). The pull of mimetism is so compelling that even adults tend to accommodate their speaking rate, intensity, and other characteristics of their speech to their interlocutor's (See the review of the speech imitation literature in Honorof, Weihing & Fowler, 2011 and in Rojczyk et al., 2013). It is no wonder, then, that the "listen and repeat" traditional approach to pronunciation teaching and learning (Jones, 1997) is also based on imitation. Learners are invited to listen to models speaking the L2 and imitate them as they repeat what they heard.

A variation of this approach has learners base their practice not on the L2 but on their L1. In this approach (Everitt, 2015), learners practice their L2 pronunciation by speaking their L1 with an L2 accent. Esling and Wong (1983) highly recommend this practice, as it leads learners to build awareness of the L2's voice quality settings (i.e., the position of the larynx, pharynx, tongue, etc. typical of the L2). They also suggest learners imitate an L2 accent in their L1 using a stereotypical L2 accent, as stereotypical accents "often include visible characteristic vocal postures" of a language (1983: 94).

Stereotypes are often wrapped in a negative light and subsequently rejected because of the fragmentary and reductive perspective they offer. Stereotypical accents, which are built from a selection of the segmental and suprasegmental features of an authentic accent (Kristiansen, 2001), are usually avoided in second language acquisition for similar reasons. However, they present some positive characteristics. Their exaggerated nature makes their features perceptually salient (Kristiansen, 2003), and more likely to be noticed and acquired (Schmidt, 1990, 1993). Furthermore, they are carried by the imitator's native language. The exclusive focus on pronunciation this configuration allows increases pronunciation processing ease for the learners, who do not have to additionally attend to meaning, vocabulary, grammar, etc. Finally, stereotypical accents are generally familiar to learners, who have been exposed to them from a young age through media (Lippi-Green, 1997), including movie and cartoon characters such as The Merovingian in *The Matrix Reloaded* (Silver, Wachowski, & Wachowski, 2003) and Warner Brothers' Pepé Le Pew. There is consequently a wealth of pre-existing knowledge which stereotypical accents allow to tap into, and as early as the onset of L2 acquisition (Everitt, 2015).

L2 accent imitation has been used to investigate phonological awareness both in perception and production to determine which pronunciation features were already and not yet acquired. The majority of the research examines imitation of voice onset time (VOT) of L2 stop consonants and reports general success in reproducing the feature in both the L2 and in the imitated L2 accent in the L1 (Flege & Hammond, 1982; Mora et al., 2014; Neuhauser, 2011; Sypiańska & Olender, 2016).

Flege and Hammond (1982) examined the pronunciation of L1 English beginner learners of L2 Spanish familiar with Spanish-accented English. They were interested in learners' awareness of two Spanish phonetic characteristics in contrast with English: the shorter duration of syllables in utterance-final position and the short-lag VOT (i.e., lack of aspiration) of voiceless stop consonants. Learners were successful in producing target-like syllable duration and stop consonant VOTs in both Spanish-accented English and in Spanish. The authors conclude that

since such differences are only phonetic in nature, they shouldn't represent an obstacle to VOT reduction in L2 Spanish.

While Flege and Hammond were concerned with VOT reduction from English to Spanish, Mora and his colleagues (2014) were interested in its increase from Spanish (L1) to English (L2). Learners produced significantly higher VOTs for English voiceless stops in English and English-accented Spanish than in Spanish, suggesting awareness of the L1-L2 VOT contrast. Results also showed that they were able to modify their VOTs in English-accented Spanish to the extent that they could modify them in English, indicating that the extent of the learners' awareness of this L1-L2 contrast is related to their L2 phonological development. In the present study, participants were expected to be already aware of the L1-L2 <*r*> contrast after receiving approximately 180 hours of instruction in French. However, although it is only phonetic in nature, as was the VOT contrast in Flege & Hammond (1982), French /ʁ/ is notoriously difficult to pronounce, as it involves an articulatory configuration—drawing *back* the tongue to form a pharyngeal, velar, or uvular constriction (Tranel, 1987)—that is absent from the English repertoire. The focus of this study, then, is on the production of French /ʁ/, and on the impact on its development that explicit instruction and practice through L2 accent imitation may have.

Research on L2 accent imitation in the L1 as an aid for L2 pronunciation development is scant. Everitt (2015), who investigated the impact of this aid with L1 Spanish/Catalan learners of L2 English, studied perception and production of L2 English voiced (/b/) and voiceless (/p/, /t/ /k/) word-initial stops. Although learners receiving English accent training in Spanish did not significantly outperform learners receiving English accent training in English on perceiving English voiced and voiceless stops /b/ and /p/, they produced more target-like English voiceless stops /p/, /t/, and /k/. Everitt's results strongly suggest that accent imitation is an efficient tool for L2 pronunciation improvement.

The present study is part of a larger one designed to contribute to knowledge about the benefits of L2 accent imitation in the L1 as a pronunciation learning tool. In this larger study, several characteristics of French pronunciation and their relation to accentedness, comprehensibility, and intelligibility are examined: short-lag VOT of voiceless plosives as in the research mentioned above, but also French /ʁ/, vowel stability (no diphthongization and no vowel reduction to schwa in unaccented syllables), the front vowels /ø/ (as in *peux*, may) and /y/ (as in *sud*, south), and intonation (rise/fall) on the last syllable of the accentual phrase. They are selected from an inventory of French pronunciation features that are typically difficult for L2 learners identified by Walz (1980). In the present study, results related to /ʁ/ and accentedness are reported.

Research Questions

Three groups were created to examine the research questions. Training was based on the use of: stereotypical French accent in English (Group S); authentic (i.e., non-exaggerated) French accent in English (Group A), and French accent in French, the traditional approach (Group F).

1) Does L2 accent training in the L1 favor production development of French /ʁ/ more significantly than L2 accent training in the L2? In other words, will Group S and Group A outperform Group F?

2) Does stereotypical (i.e., exaggerated) L2 accent training in the L1 favor production development of French /ʁ/ more significantly than authentic (i.e., not exaggerated) L2 accent training in the L1? That is, will Group S outperform Group A?

METHODS

Participants

Eleven intermediate French students enrolled in a French pronunciation course at an American university and five native speakers of French participated in this study. As summarized in Table 1, eight students were assigned to each of two experimental groups, and three to a control group. The production of two native speakers served as reference for the rating process, which was completed by the remaining three native speakers.

Table 1

Participants in the three groups

Experimental Groups		**Control Group**
Stereotypical Accent Group (Group S) n=4	Authentic Accent Group (Group A) n=4	French Group (Group F) n=3

Stimuli

All three groups received explicit instruction on the characteristics of French accent. As data collection is on-going, this paper reports on results associated with /ʁ/. Participants were instructed on the articulatory differences between American English and French *r*. Tranel (1987) explains that the degree of frication and voicing of French /ʁ/ varies partly according to its position in the utterance and in the syllable (i.e., its adjacency to voiced and voiceless sounds). Colantoni and Steele (2007) found that English learners of L2 French master manner before voicing, in salient position (CV, e.g., *Pa`ris*) before less salient one (VC, e.g., *sûr*, sure) due to hyperarticulation leading to the production of overly long—and consequently—devoiced /ʁ/. The participants in this study are expected to follow this pattern and receive higher scores for words featuring /ʁ/ in CV context as displayed in the left column of Table 2.

Table 2

/ʁ/ stimuli

CV n=4	**VC** n=5
Paris	*Bonjour*, hello
Entrez, come in	*Heures*, hours
Française, French	*Journée*, day
Trente, thirty	*Martin* (name)
	Sûr, sure

Training

Each group additionally underwent one weekly 20-minute session over the course of three weeks. During the session, they received further instruction about the French accent characteristics as they were featured in the speech of the models analyzed in class, which participants were to imitate during pronunciation practice at home. In Group S, the models were non-native speakers of French speaking English with a stereotypical French accent. In Group A, they were native speakers of French and spoke English with an authentic (i.e., non-stereotypical, unexaggerated) French accent. And the students in Group F based their pronunciation practice on the traditional approach by modeling their speech on that of French native speaker models speaking French. Care was taken to feature a different model for each session, as high variability phonetic training has been shown to enhance pronunciation acquisition (Lively, Logan & Pisoni, 1993; Bradlow, Pisoni, Akahane-Yamada & Tohkura 1997; Thomson, 2011).

Tests

Students were instructed to practice imitating their models at home following each session, and record their best imitation and submit it to their instructor (the author). Additionally, they recorded themselves three times: before and immediately after treatment, and a week later to measure for long-term effects. For each test, they read a French narrative and a dialogue (the same for each test), described a picture, and created a dialogue (new ones for each test). The present study reports results related to accentedness and /ʁ/ in the read narrative and original dialogue.

Rating procedure

The nine words featuring /ʁ/ were extracted from the 11 participants' recorded narrative and dialogue before, immediately, and one week after treatment. The same words from the native French speakers were added for reference. Four recordings had to be discarded due to poor acoustic quality, making them impossible to rate. A total of 347 tokens were submitted for rating to three French native speakers familiar with the speech patterns of English learners of L2 French. They each rated all tokens using a nine-point Likert-type scale (Derwing, Rossiter, Munro & Thomson, 2004) ranging from 1—"Very strong accent" to 9—"No accent."

RESULTS

The degree of agreement between judges was calculated and found to be high: the average measures intraclass correlation coefficient was .918 with a 95% confidence interval from .901 to .932 ($F(346,692) = 12.131$, $p < .001$). A one-way ANOVA revealed no significant differences between the three groups in their production of /ʁ/ at the pre-test, indicating that the proficiency level of the groups was similar before treatment ($F(2,8) = .136$, $p = .875$).

A repeated measures ANOVA with Group as a between-subjects factor and Time as a within-subjects factor was performed for subjects' average ratings on all words featuring /ʁ/. Results displayed in Table 3 indicate no significant pronunciation improvement of /ʁ/ and no significant difference between groups immediately and one week after treatment.

Table 3

Tests of Within- and Between-subjects Effects from the Repeated-Measures ANOVA (All Words with /ʁ/ combined)

Source	Type II Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	.293	2	.147	.499	.616	.059
Group	2.803	2	1.402	.630	.557	.136
Time * Group	.668	4	.167	.568	.689	.124

In fact, group means would suggest that in general, practice with models speaking French is more beneficial—although not significantly so—than practice with a model speaking English, be it with or without a stereotypical accent, as illustrated in Figure 1.

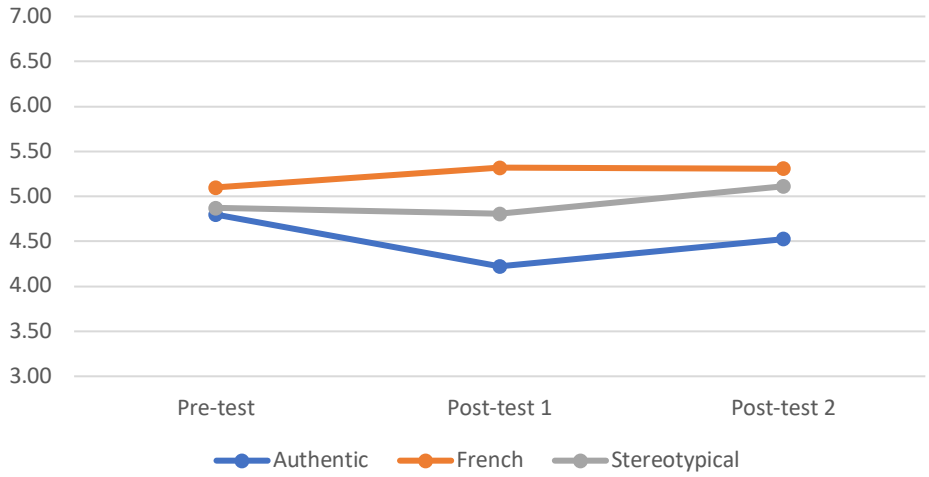


Figure 1. Group Means, Pre-test, Immediate and Delayed Post-test (All Words in /ʁ/ combined)

When breaking down the analysis by feature as in Table 4, i.e., when distinguishing /ʁ/ in CV environments from /ʁ/ in VC environments, the effects of treatment and time are still negligible.

Table 4

Tests of Within-subjects Effects from the Repeated-Measures ANOVA (Feature = CV and VC)

Source	Type II Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Time	.532	2	.266	.498	.617	.059
Time * Group	1.565	4	.391	.733	.582	.155
Feature	34.229	1	34.229	24.513	.001	.754
Feature * Group	.275	2	.138	.099	.907	.024
Time * Feature	2.379	2	1.190	2.637	.102	.248
Time * Feature * Group	1.030	4	.258	.571	.688	.125

Results for "Feature" ($F(1, 34.23) = 24.51$, $p < 0.001$, partial $\eta^2 = 0.75$), as well as Figure 2 below do indicate that regardless of practice type, all participants obtained significantly higher scores for CV than for VC environments at all times. These results are in line with those of Colantoni and Steele (2007) and confirm expectations that L2 learners of French improve their pronunciation of /ʁ/ appearing in the onset of a syllable (e.g., *Pa`ris*) before they do that of /ʁ/ in coda position (e.g., *sûr*).

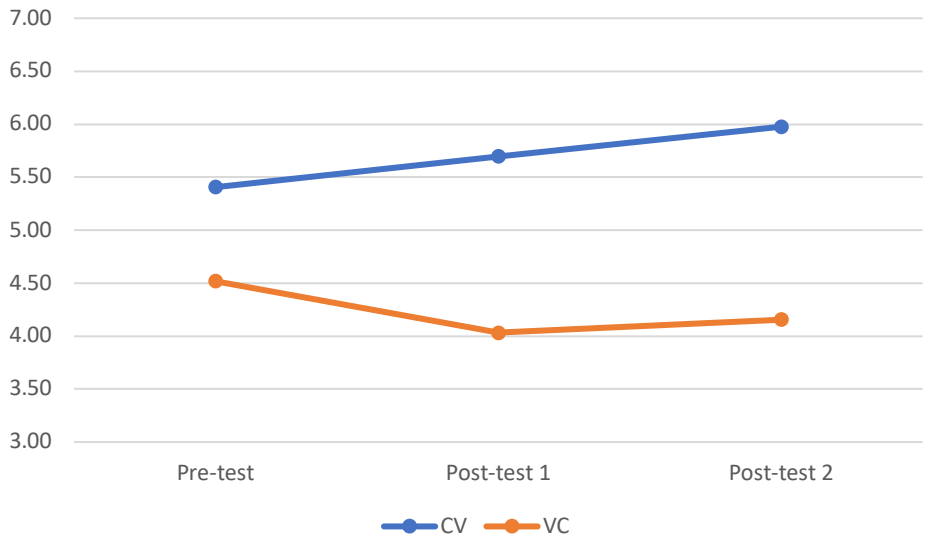


Figure 2. /ʁ/ Feature (CV and VC) Means from Pre-test to Immediate and Delayed Post-test

DISCUSSION

The absence of a significant difference between groups immediately and one week after treatment suggests that for the participants in this study, the use of stereotypical accent for pronunciation improvement of French /ʁ/ was not effective. Indeed, salience of—and familiarity with— the stereotypical version of this feature did not give subjects in the Stereotypical Group an advantage at improving their pronunciation of /ʁ/. Nor did alleviated processing of this feature, made possible by exclusive concentration on pronunciation, which could have favored both the Stereotypical and the Authentic Groups.

It is to be noted, however, that the lack of improvement on both kinds of /ʁ/ in the Authentic Group is not surprising, as only one of the three models imitated in this condition pronounced a French /ʁ/, leaving participants with considerably less practice of the feature. This makes for one of the limitations of this study, and treatment applied to additional features of French pronunciation, such as vowel stability and intonation, may yield positive results, including improved general intelligibility and comprehensibility.

Other limitations include a small number of participants, and perhaps insufficient time on task, as well as focus on too many features at a time. No significant improvement after treatment for any condition suggests that more practice may be needed. In a follow-up survey, all participants indicated that they would welcome more instruction along those lines, but over half of them (6/11, 2 in each of the three groups) explicitly suggested more training and practice. Moreover, one respondent expressed difficulty at processing several pronunciation features at once, and all of the other participants mentioned that they were grateful for the notes they had taken during training. This suggests that future treatment limiting the focus of each training session to the study of one to two pronunciation features may make for easier processing.

Finally, the results may lead us to question the popularity of this approach. Over half of the respondents who practiced with a model speaking English (6/8 respondents) said they enjoyed doing so, as it allowed them to focus solely on pronunciation. However, four of them (two in each of the Authentic and the Stereotypical Groups) reported they would have preferred practicing with French native speaker models speaking French. This sentiment is to be expected from learners enrolled in a pronunciation course, who would themselves expect practice based on models speaking French, and may not reflect the reaction of learners in courses not specifically focusing on pronunciation. Further research is, therefore, needed before it is to be considered inefficient for French pronunciation improvement.

ACKNOWLEDGEMENTS

I am grateful to Cinthia Medina Mendoza and Jesse Klaus, my research assistants, and to three native speaker raters: Véronique Jewell, Sylvie Lidolf, and Violène Pillar. This work was supported by grants from the College Dissemination and Discovery Award and the Support for Faculty Scholars Award, Western Michigan University.

ABOUT THE AUTHOR

Viviane Ruellot is an Associate Professor of French at Western Michigan University, where she teaches French, French linguistics and applied linguistics. Her research interests include

feedback on pronunciation, the stages of pronunciation acquisition, computer assisted pronunciation teaching and learning, and the history of French pronunciation teaching.

Department of World Languages and Literatures
821 Sprau Tower
1903 W. Michigan Ave.
Kalamazoo, MI 49008-5538

(269)387-3043

viviane.ruellot@wmich.edu

REFERENCES

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101* (4), 2299-2310.

Chun, D. (1991). The state of the art in teaching pronunciation. In J. Alatis (Ed.), *Linguistics and Language pedagogy: The state of the art. Georgetown University Round Table on Languages and Linguistics* (pp. 179–193). Washington, DC: Georgetown University Press.

Colantoni, L., & Steele, J. (2007). ACQUIRING /[alveolar approximant]/ IN CONTEXT. *Studies in Second Language Acquisition*, *29*(3), 381-406.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. Amsterdam, the Netherlands: John Benjamins.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language learning*, *54*(4), 655-679.

Esling, J. H., & Wong, R. F. (1983). Voice quality settings and the teaching of pronunciation. *TESOL quarterly*, *17*(1), 89-95.

Everitt, C. (2015). *Accent Imitation on the L1 as a task to improve L2 pronunciation*. Unpublished MA thesis. University of Barcelona: Barcelona, Spain

Flege, J. E., & Hammond, R. M. (1982). Mimicry of non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition*, *5*(1), 1-17.

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

Honorof, D. N., Weihing, J., & Fowler, C. A. (2011). Articulatory events are imitated under rapid shadowing. *Journal of Phonetics*, *39*(1), 18-38.

Jones, R. H. (1997). Beyond “listen and repeat”: Pronunciation teaching materials and theories of second language acquisition. *System*, *25*(1), 103-112.

Kristiansen, G. (2001). Social and linguistic stereotyping: A cognitive approach to accents. *Estudios ingleses de la Universidad Complutense*, (9), 129.

Kristiansen, G. (2003). How to do things with allophones: Linguistic stereotypes as cognitive reference points in social cognition. *Cognitive Models in Language and Thought:*

Ideologies, Metaphors, and Meanings. Berlin/New York: Mouton de Gruyter, 69-120.

Kuhl, P. K., & Meltzoff, A. N. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *The journal of the Acoustical Society of America*, *100*(4), 2425-2438.

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics, 36*(3), 345–366.

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

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*(3), 1242-1255.

Mora, J. C., Rochdi, Y., & Kivistö-de Souza, H. (2014). Mimicking accented speech as L2 phonological awareness. *Language Awareness*, *23*(1-2), 57-75.

Nagell, K., Olguin, R. S., & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees (Pan troglodytes) and human children (Homo sapiens). *Journal of Comparative Psychology*, *107*(2), 174.

Neuhauser, S. (2011). Foreign accent imitation and variation of VOT and voicing in plosives. In *Proceedings of the XVIIth International Congress of Phonetic Science, Hong Kong* (pp. 1462-1465).

Rojczyk, A., Porzuczek, A., & Bergier, M. (2013). Immediate and distracted imitation in second-language speech: Unreleased plosives in English. *Research in Language*, *11*(1), 3-18.

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *Tesol Quarterly*, *46*(4), 842-854.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied linguistics*, *11*(2), 129-158.

Schmidt, R. W. (1993). Awareness and second language acquisition. *Annual review of applied linguistics*, *13*, 206-226.

Silver, J., (Producer) & Wachowski, L. & Wachowski, L., (Directors). (2003). *The Matrix Reloaded* [Motion picture]. United States: Warner Home Video.

Sypiańska, J., & Olender, A. (2016). *Phonetic awareness in Polish learners of English* (pp. 1-9). Poznań Working Papers in Linguistics, *1*(1).

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *Calico Journal, 28*(3), 744-765.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics, 36*(3), 326-344.

Tranel, B. (1987). *The sounds of French: An introduction*. Cambridge university press.

Walz, J. (1980). An empirical study of pronunciation errors in French. *The French Review*, *53*(3), 424-432.

Zentall, T. & Akins, C. (2001). Imitation in animals: Evidence, function and mechanisms. In R. G. Cook (Ed.), *Avian visual cognition*. [On-line]. Available: www.pigeon.psy.tufts.edu/avc/zentall/.

Mora-Plaza, I. Mora, J. C., & Gilabert, R. (2018). Learning L2 pronunciation through communicative tasks. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 174-184). Ames, IA: Iowa State University.

LEARNING L2 PRONUNCIATION THROUGH COMMUNICATIVE TASKS

Ingrid Mora-Plaza, University of Barcelona

Joan C. Mora, University of Barcelona

Roger Gilabert, University of Barcelona

L2 pronunciation is often neglected in the EFL classroom and, when addressed, it is typically decontextualized from communicative practice. Additionally, limited research has been conducted in SLA on the role of task manipulation for the improvement of L2 pronunciation accuracy during meaning-focused interaction. This study investigated the impact of decision-making tasks, organized in increasing complexity, on the production of English /æ/-/ʌ/. L1 Catalan/Spanish young adults (*n*=18) performed four dyadic problem-solving, reasoning-gap tasks over a three-week period. Tasks were always preceded by form-focused pre-tasks that contained lexical items contrasting the target vowels (e.g., *bag-bug, cap-cup*) to be used during task performance. Furthermore, tasks were sequenced on the basis of increasing level of cognitive complexity (+S, -S, -C, +C) in order to progressively enhance the occurrence of pronunciation-based language-related episodes (LRE). Production accuracy was pre- and post-tested through a delayed-sentence repetition task. In line with the Cognition Hypothesis (Robinson, 2007, 2011), the results revealed that orienting attention to a phonological contrast during interactive tasks improves its production significantly, and increased task demands along resource-directing variables (i.e. ± reasoning demands and ± elements) generate more pronunciation-focused LREs.

INTRODUCTION

The learning of second language (L2) phonological representations requires practice through long periods of exposure to the foreign language (FL). Nevertheless, in school contexts, authentic L2 oral input may be scarce and often limited in out-of-class exposure (Muñoz, 2008). Apart from the lack of linguistic experience in the FL environment, L2 pronunciation is conceived as one of the most challenging skills to be taught and learned in English as a FL classrooms. According to Murphy and Baker's (2015) historical overview of the teaching of pronunciation, none of the current methodologies appears to be effective enough due to an interplay of factors such as the use of old methods and outdated materials, and lack of teacher training. Despite this, Guion-Anderson and Pederson (2007) claimed that pronunciation-based instruction is likely to help learners 'notice the gap' between L1 and L2 phonetic categories and produce more accurate pronunciation.

Nowadays, many practitioners follow an analytic focus-on-form approach, which implicitly draws learners' attention to form in the context of meaningful communication. This approach is motivated by the Interaction Hypothesis (Gass, 1997), which claims that interaction is crucial in L2 acquisition, and the modifications that result from negotiation of meaning increase input and output comprehensibility. In any case, phonetic learning is not instantaneous and learners may first exhibit emergent interlanguage forms that need to be repetitively practiced in content-based contexts in order to be internalised (Saito, 2013). Saito and Wu (2014) advocate for orienting attention to phonetic form while maintaining the primary focus on meaning, and they emphasize the benefits of integrating instruction on

suprasegmental features in formal teaching environments. As negotiation of form in content-based lessons has been shown to improve L2 pronunciation accuracy, tasks are a useful instrument to direct learners' cognitive resources to phonetic forms during real-world activities (Salaberry & López-Ortega, 1998).

From a task-based language teaching (TBLT) perspective, tasks ought to have a clear goal and well-defined outcome that learners need to fulfil. Also, tasks increase the conditions for focus on form during communicative activities that bear resemblance to real-world tasks and involve cognitive processes that promote L2 development and performance. Additionally, task features can be manipulated in order to generate further focus on form in meaning-driven interactions. This study follows the Cognition Hypothesis (Robinson, 2007, 2011) which emphasizes the flexibility of attentional capacity and claims that greater effort at conceptualization induces learners to stretch and develop their L2 linguistic resources. Robinson and Gilabert (2007, p. 162), following the SSARC[1] model, state that "pedagogic tasks should be designed, and then sequenced for learners on the basis of increases in their cognitive complexity" because such tasks have the potential to lead to more accurate and complex language. Within resource-directing dimensions, tasks can be manipulated by increasing task complexity through ± elements and ± reasoning demands, each of which guides resources to specific functional and linguistic requirements (Talmy, 2000).

For example, Gilabert, Baron & Llanes (2009) and Baralt (2014) found that complex tasks where learners have to exploit their attention and memory resources trigger more language-related episodes (LREs). These can be defined as any part of a dialogue where learners discuss language they are producing, question their language use, or self-correct their language production (Swain & Lapkin, 1995). Our study considered self-repairs (i.e., the learners' self-correction of faulty pronunciation), recasts (i.e., a correct restatement of a learner's incorrectly formed utterance) and repetitions (i.e., the learners' statement of the same word with the same pronunciation) as instances of LREs. Increasing task demands along resource-directing dimensions is likely to draw attention to how messages are being encoded during performance and, consequently, lead to interlanguage development (Gilabert, 2007). Nevertheless, complex tasks are more prone to inducing LREs when they are not extremely challenging and understanding between interlocutors is sufficient for communication (Révész, 2011).

Task-based pronunciation teaching (TB*P*T) presents tasks which generate form-focused episodes that target phonological elements during interaction. In other words, tasks raise awareness of pronunciation elements by making target items essential and enhancing the occurrence of pronunciation-focused LREs during interaction (Mora & Levkina, 2017). Empirical studies on TB*P*T are limited, but some researchers have successfully applied already extensively researched TBLT dimensions (i.e., task complexity, task repetition and task modality) to L2 pronunciation. One example is Solon, Long, & Gurzynski-Weiss (2017), whose findings support the Cognition Hypothesis (Robinson, 2007, 2011) in that the more complex version of the task generated more accurate realizations of L2 vowels; however, contrary to what might be expected, the simple task generated more pronunciation-based LREs than the complex task.

[1] SSARC stands for *stabilize, simplify, automatize, restructure*, and *complexify*. This model posits that (a) task sequencing should be based on cognitive complexity factors and (b) tasks should increase first in resource-dispersing dimensions and then, in resource-directing dimensions.

THE PRESENT STUDY

The aim of this study is to bridge the gap between L2 pronunciation instruction and task-based language teaching. The design of this experiment was based on four communicative tasks, embedded in a real-world situation, targeting L2 phonological forms that were essential for task completion. Specifically, the goal was to examine the effectiveness of task design on the production of a difficult vowel contrast for EFL learners. The selected phonological contrast was /æ/-/ʌ/ (e.g. *cat-cut*), two English sounds that are known to be challenging for Catalan/Spanish speakers because they are perceptually assimilated to a single L1 low vowel category /a/ (Flege, 1995; Best & Tyler, 2007; Rallo-Fabra & Romero, 2012). Task complexity was operationalized as manipulation of cognitive complexity differences between 4 tasks in terms of ± elements and ± reasoning demands. The objective was to enhance pronunciation-based LREs, which would facilitate the improvement of L2 pronunciation accuracy. Considering the objectives outlined above, our study addressed the following research question:

> Does increasing task complexity have an effect on the occurrence of pronunciation-based LREs? Is the frequency of pronunciation-based LREs related to learners' gains in the production of the /æ/-/ʌ/ contrast?

METHODS

Participants

Thirty-six Catalan-Spanish bilingual EFL learners (18 controls) at secondary school took part in the study. In the experimental group, there were 18 students (9 females, mean age 16.4) learning English together at school since the age of 6. 61.1% of the class stated that they had received extra-curricular English instruction between 2 and 13 years (M= 6.36) for 2.5h/week. The vocabulary size test (M= 2933.33) indicated a B1-B2 level according to the Common European Framework of Reference for Languages. Other foreign languages spoken were French (B1) [5 students], Italian (A2) [1 student] and Portuguese (A1) [1 student].

Instruments

The experimental and control groups were tested before and after a one-week treatment period, which involved 4 daily sessions of 15 to 30 minutes each. Production was tested through a delayed sentence repetition task. The pre-test and post-test were identical except for some novel items included in the latter in order to assess generalization of gains. Novel items were non-words that learners had not been exposed to during testing or training, or items learners had been previously exposed to, but produced by a different voice. The target items were practiced repetitively during pre-task and tasks.

Stimuli

The selected target sounds were the two standard Southern British English vowels /æ/ and /ʌ/. This contrast was embedded in 10 pairs of real words, which appeared in the pre-test and post-test, and 5 pairs of non-words, which appeared only in the post-test. They were uttered by four English native speakers (two males).

Testing

In the delayed sentence repetition task, learners were exposed to 44 test trials + 2 practice trials. The 44 test trials included 40 sentences with the target vowels /æ/ and /ʌ/ (e.g., Your

cap is on my head. A *cup* of tea, please.) and 4 sentences with other non-target minimal pairs such as /iː/ and /ɪ/ (e.g., The *sheep* are eating flowers. The *ship* is alone in the sea.) These distractors were used in order to get a more natural exposure by avoiding learners focusing too much on the target vowels. The sentences were uttered once by one male and one female speaker.

Training

Pre-tasks

A general pre-task was used to train the meaning of the words that appeared in the tasks. Then, participants practiced their pronunciation through word imitation and sentence imitation tasks where feedback was provided. In the word imitation and sentence imitation, 10 minimal pairs were presented (20 tokens), containing the vowels /æ/ and /ʌ/ (i.e. *bag/bug, bat/butt, cap/cup, cat/cut, mag/mug, ram/rum, amber/umber, babble/bubble, natty/nutty, stab/stub*) plus 5 minimal pair distractors (10 tokens), containing the vowels /iː/ and /ɪ/ (i.e. *bean/bin, feast/fist, sheep/ship, teen/tin, weep/whip*) with their corresponding images.

Tasks

The four decision-making tasks were two-way, split, close and convergent (Pica, Kanagy & Falodun, 1993) because the two interlocutors had different information and had to come up with one single solution (Figure 1). Moreover, learners could not solve the task without producing the L2 phonological contrast accurately and so, the contrast was made 'task-essential' language (Loschky & Bley-Vroman, 1993). These tasks were designed around a trip to Kenya that students had to plan. In a sequential manner, learners had to decide on what they wanted to see and buy in a natural park in Kenya (Task 1); the objects they wanted to bring to Kenya (Task 2); the organization of a roleplay party in Kenya (Task 3) and what they wished to post on the school website (Task 4). All tasks involved two mental operations: information-sharing and decision-making. Nevertheless, in order to complexify the tasks in increasing order (+S, -S, -C, +C)[2], more elements and reasoning demands were added over the course of a week (see SSARC model, Robinson, 2010). In addition, task complexity was independently assessed (Révész, 2011) by ten experienced language teachers, who critically evaluated the tasks according to degree of difficulty and mental effort. Students themselves also rated task difficulty after each task. All the tasks had been previously piloted on a similar sample of learners and proved to be adequate in difficulty for their level of proficiency.

[2] S= Simple
C= Complex

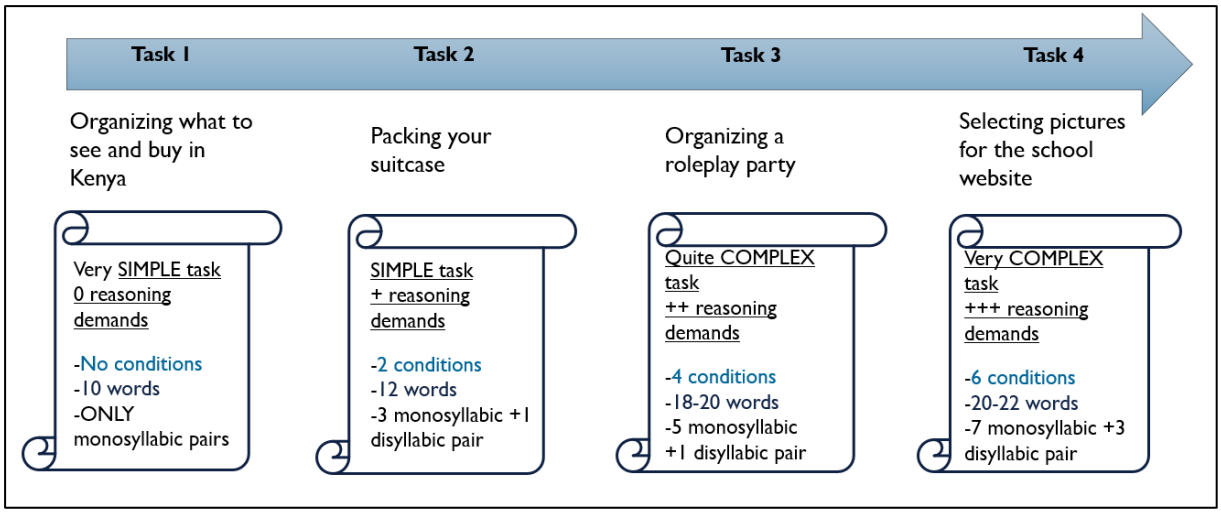


Figure 1. Sequential organization of the tasks in order of increasing complexity.

Procedure

In the DSR task, learners had to read the sentence on the screen, listen to it, wait for a beep sound and repeat it. Learners' productions were recorded with a digital *Tascam Dr-40* recorder with an external *Shure SM58* microphone. After the pre-test, learners in the experimental group did a pre-task and several tasks. Tasks lasted 15-30 minutes, depending on their complexity and were registered using two *Tascam Dr-40* solid-state recorders. Learners were placed in front of a microphone and facing each other so they could not see each other's piece of information.

Analysis

Learners' accuracy scores were obtained by analysing the quality of vowels /æ/ and /ʌ/. Analyses focused on the first formant (height) and second formant (advancement), which were transformed into Bark following Syrdal & Gopal's (1986) formula (B_i = 26.81/(1+1960/F_i) - 0.53) and then, the spectral distance between the two vowels was also calculated.

Three raters (mean age 24.7), who were experienced English teachers living in Barcelona (Spain) at the time of testing, were instructed on the analysis of LREs. They listened to all the recordings (4 tasks), transcribed the pronunciation focused LREs and classified them into four types: (a) general LREs, (b) recasts, (c) self-repairs, and (d) repetitions. The number of LREs per person was calculated with all types of LREs as well as with only general LREs. In addition, an LRE ratio (LRE/time-on-task) was estimated to interpret the results in accordance with Solon et al. (2017). It was important to compensate for the different lengths of the productions, generated by the different levels of complexity, with more complex tasks generating more interaction. Finally, in order to correlate production gains and number of LREs, each pair of students was assigned the same number of LREs, irrespective of which student in the dyad initiated or produced the LREs.

RESULTS

Total inter-rater reliability across tasks was 91.6%: 76.7% (task 1), 97.8% (task 2), 95.2% (task 3) and 96.6% (task 4). A one-way ANOVA was used to assess the effect of task complexity (+S, -S, -C, +C) on the occurrence of all LREs. The ANOVA revealed a significant main effect of task complexity, (F (3,15) = 42.630, p<.001, η^2 =.895). Bonferroni-

adjusted pairwise comparisons showed that the number of LREs increased significantly as cognitive complexity increased across tasks (Figure 2): task 1 (*M*=3.89, *SD*=2.96), task 2 (*M*=6.00, *SD*=2.91) and task 3 (*M*=10.00, *SD*=4.82), (*p*<.05). However, the difference in the number of LREs between task 3 and task 4 did not reach significance (*M*=13.11, *SD*=4.56, *p*=.086).

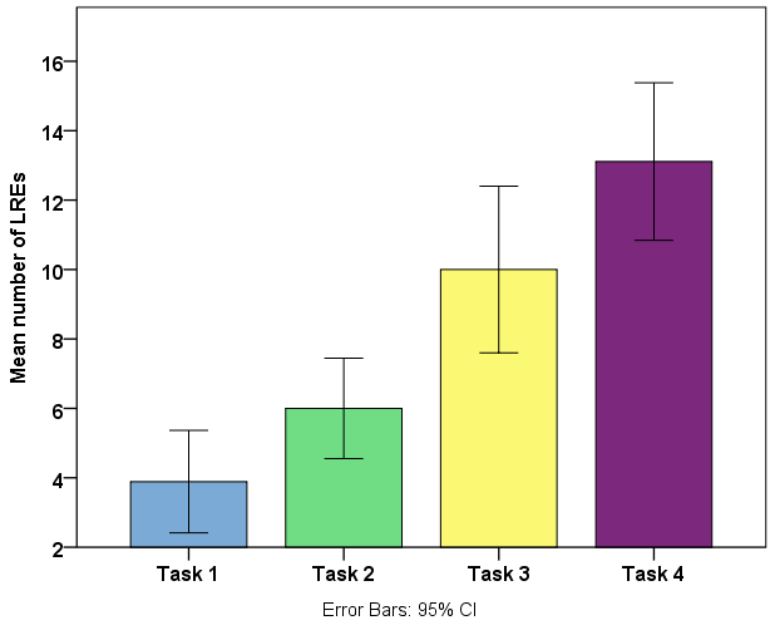


Figure 2. Mean number of LREs (general LRE, recasts, self-repairs and repetitions) by task.

When analysing LREs per minute, the one-way ANOVA confirmed a significantly lower number of LREs in simple than in complex tasks, (F (3,15) = 7.747, p=.002, η^2 =.608) (Figure 3), a cognitive complexity effect mainly driven by the significant difference between task 1 and tasks 2, 3, 4 (p<.05), as differences between tasks 2, 3 or 4 did not reach significance. In short, the more complex the task was, the higher the occurrence of pronunciation-based LREs irrespective of time-on-task.

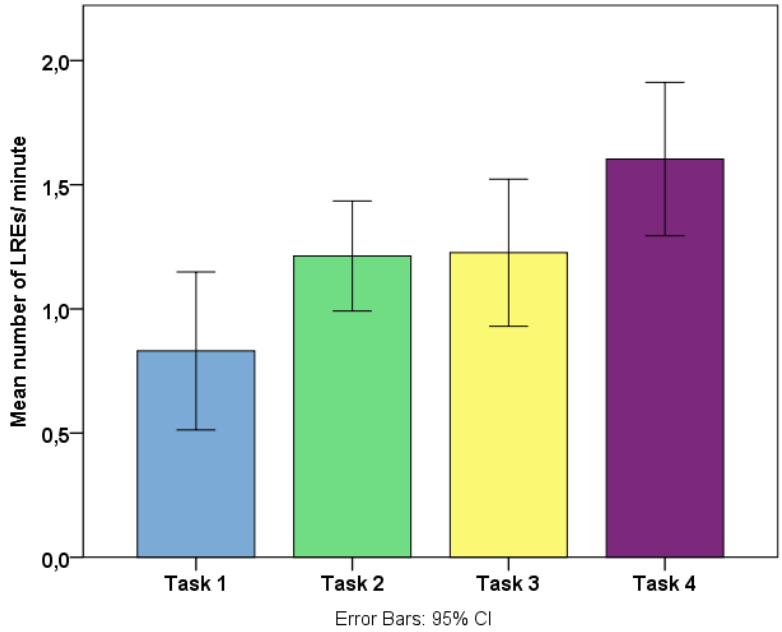


Figure 3. Mean number of LREs per minute by task.

Production results showed that learners' Euclidean distance increased (albeit non-significantly) from pre-test to post-test (Table 1) owing to the treatment with tasks. LREs were analysed in comparison to learners' accuracy gains in the production of /æ/ and /ʌ/. The *Pearson-r* correlation revealed that the more language related episodes learners produced, the larger the size of gains in Euclidean distance learners obtained (r=.479, p=.044). As a result, we can conclude that the more often they paid attention to the phonological contrast, the better able they were to distinguish between the two vowels in production.

Table 1

Descriptive statistics for Euclidean distances between /æ/ and /ʌ/.

	Pre-test			Post-test		
	N	*M*	*SD*	*N*	*M*	*SD*
Males	9	1.45	.65	9	1.47	.64
Females	9	1.45	.45	9	1.54	.66

DISCUSSION AND CONCLUSIONS

According to the Cognition Hypothesis (Robinson, 2007, 2011), greater task complexity enhances greater incidence of form-focused episodes, which would promote interlanguage development. Following previous studies on grammar (Baralt, 2013) and pragmatics (Kim & Taguchi, 2015), our study has shown that four decision-making tasks, with a focus on pronunciation, improved learners' L2 pronunciation accuracy. Furthermore, in line with Baralt (2013) and Gilabert's (2007) findings, learners' attention towards language increased as tasks increased in complexity because more complex tasks demanded the use of more

precise linguistic resources. In addition, the reduction of resource-dispersing variables (i.e., familiarity with the task and planning time) helped them pay attention to phonetic form during communicative interaction.

The results of the present study show that a significantly larger number of pronunciation-based LREs was generated in more complex tasks. Interestingly, the occurrence of LREs per minute was also higher in the complex tasks, which indicates that learners were reflecting on form to a higher extent in complex tasks regardless of time-on-task. This finding is at odds with Solon et al.'s (2017) findings, which found that simpler tasks generated more LREs (albeit not significantly). They argued that whereas grammatical targets have specific forms that can be described by metalinguistic rules, phonetic targets cannot be described through such rules because they are part of the gradient range of production possibilities and require the physical modification of the articulators. In addition, learners were not familiar with verbally reflecting on phonetic form as pronunciation training was not a part of their language curriculum. However, we interpret our results to suggest that it is possible to raise metaphonological awareness by generating a clear focus on phonetic form and by making the target phonological contrast essential for task completion during meaningful interaction. Under such circumstances, learners appear to be able to negotiate the target form explicitly and implicitly through direct corrections, recasts or repetitions. As in Solon et al. (2017), pronunciation training was not a part of the learners' curriculum and, even if they were not accustomed to reflecting verbally on phonetic form, learners developed strategies to improve intelligibility during conversation. The result was a higher production of LREs regardless of the real time-on-task.

The present study has also shown that the number of LREs was related to the size of gains in production. The careful task design triggered many opportunities to focus on the language learners were producing (cf. Sicola, 2009), especially those phonetic forms implementing the target phonological contrast.

To conclude, our study has contributed to current research investigating whether the benefits of tasks can be extended beyond grammar and lexis to L2 pronunciation (Gurzynski-Weiss, Long, & Solon, 2017) and how task manipulation can help enhance an intentional focus on phonetic form. Through the use of four carefully-designed pedagogic tasks that resemble real-world events, we called learners' attention to phonetic form during meaningful interaction. In this way, learners were able to notice the gap between their peers' productions and their own as well as engage in metaphonological reflection on the phonetic form of the output (Robinson, 2011). Such a method respected learners' developmental stages and processing ability while making the target form essential for task completion. Increasing cognitive complexity through task design resulted in higher occurrence of LREs, which led to improvement in L2 segmental accuracy. Our study provides evidence of the potential benefits of task design and manipulation on L2 pronunciation development, and suggests that TB*P*T is a promising research avenue that will help bridge the gap between research and practice in L2 pronunciation teaching and learning.

ACKNOWLEDGMENTS

We would like to thank our participants, interraters and graphic designers for their excellent job as well as the audiences at the 2017 Pronunciation in Second Language Learning and Teaching Conference in Salt Lake City, Utah for their comments and feedback.

ABOUT THE AUTHORS

Ingrid Mora-Plaza is a Ph.D. candidate in the Cognitive Science and Language program and lecturer at the University of Barcelona (UB). Her primary area of specialization is at the interface between task-based language teaching (TBLT) and L2 pronunciation instruction. She also does research on the impact of individual differences in cognitive attention control (i.e., auditory selective attention and auditory inhibition) on L2 speech acquisition. Email: imoraplaza@ub.edu

Joan C. Mora is associate professor in the Department of Modern Languages and Literatures and English Studies in University of Barcelona (UB). His research has examined the role of contextual and individual factors in the development of L2 speech and oral fluency, and the acquisition of L2 phonology. Email: mora@ub.edu

Roger Gilabert is associate professor in the Department of Modern Languages and Literatures and English Studies at the University of Barcelona (UB). He has conducted research in the area of task design and task complexity and its impact on L2 performance and acquisition. Email: rogergilabert@ub.edu

REFERENCES

Baralt, M. (2013). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition*, 35, 689-725.

Baralt, M. (2014). Task complexity and task sequencing in traditional versus online language classes. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 95–122). London, UK: Bloomsbury.

Best, C. & Tyler, M. (2007). Nonnative and second language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. Munro (Eds.), Language Experience in Second Language Speech Learning: In honour of James Emil Flege (pp. 13-34). Amsterdam/Philadelphia: John Benjamins.

Flege, J. E. (1995). Second-language speech learning: theory, findings, and problems. In Strange, W. (ed)*, Speech Perception and Linguistic Experience: Issues in Cross-language research* (pp. 229-273). Timonium, MD: York Press.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Mahwah, NJ: Erlbaum.

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching, 45,* 215- 240.

Gilabert, R., Baron, J., & Llanes, M.A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during task performance. *International Review of Applied Linguistics, 47,* 367-395.

Guion-Anderson, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O-S. Bohn & M. J. Munro (Eds.) *Language experience in second language learning* (pp. 57-77). Amsterdam: John Benjamins.

Gurzynski-Weiss, L., Long A.Y. & Solon M. (2017). TBLT and L2 Pronunciation: Do the benefits of tasks extend beyond grammar and lexis? *Studies in Second Language Acquisition*, 39, 213-224.

Kim, Y. & Taguchi, N. (2016). Learner–learner interaction during collaborative pragmatic tasks: The role of cognitive and pragmatic task demands. *American council on the teaching of foreign languages*, 49.1, 52-57.

Loschky, L. & Bley-Vroman R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–67). Clevedon: Multilingual Matters.

Mora, J. C. & Levkina, M. (2017). Task-based pronunciation teaching and research: key issues and future directions. *Studies in Second Language Acquisition*, 39, 381-399.

Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 24(4), 578-596.

Murphy, J. & Baker, A. (2015). History of ESL pronunciation teaching. In M. Reed & J. Levis (Eds.), *The handbook of English pronunciation*, (pp. 36-65). Malden, MA: Wiley Blackwell.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language research and instruction. In G. Crookes & S. Gass (Eds.), *Tasks and second language learning* (pp. 9–34). Clevedon, UK: Multilingual Matters.

Rallo-Fabra, L. & Romero J. (2012). Native Catalan learners' perception and production of English vowels. *Journal of Phonetics*, 40, 491-508.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on speech production, interaction, uptake of premodified input and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching (IRAL), 45,* 195-215.

Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In M. Putz & L. Sicola (Eds.), *Cognitive processing in second language acquisition: Inside the learner's mind* (pp. 243-268). Amsterdam/Philadelphia, PA: John Benjamins.

Robinson, P. (2011). *Second language task complexity: researching the cognition hypothesis of language learning and performance*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176.

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: a classroom-based study. *The Modern Language Journal,* 95, 162-181.

Saito, K. (2013). Reexamining effects of form-focused instruction on L2 pronunciation development: the role of explicit information. *Studies in Second Language Acquisition,* 35, 1-29.

Saito, K. & Wu, X. (2014). Communicative focus on form and second language suprasegmental learning. *Studies in Second Language Acquisition,* 36, 647-680.

Salaberry, M. R. & López-Ortega, N. (1998). Accurate L2 production across language tasks: Focus on form, focus on meaning, and communicative control. *The Modern Language Journal,* 82.4, 514-532.

Sicola, L. (2009). "*No, they won't just sound like each other": Attention to phonological form in NNS-NNS inter action on targeted L2 pronunciation tasks*. Landau, Germany: Peter Lang Publishers.

Solon, M., Long, A.Y. & Gurzynski-Weiss, L. (2017). Task complexity, language-related episodes, and production of L2 Spanish vowels. *Studies in Second Language Acquisition, 39*, 347-380.

Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.

Syrdal, A. K. & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America, 79*, 1086-100.

Talmy, L. (2000). *Toward a Cognitive Semantics*, Vol. 1, *Concept Structuring Systems*. Cambridge, MA: MIT Press.

Tanner, M. & Chugg, A. (2018). Empowering adult ELLS' fluency and pronunciation skills through readers theater. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 185-193). Ames, IA: Iowa State University.

EMPOWERING ADULT ELLS' FLUENCY AND PRONUNCIATION SKILLS THROUGH READERS THEATER

Mark Tanner, Brigham Young University

Alisha Chugg, Brigham Young University

Readers Theater (RT) is a technique that has been used largely with elementary and middle school students (Corcoran, 2005; Keehn, Harmon, and Shoho, 2008) as a means of improving reading fluency. Few if any empirical studies have investigated the use of RT in building adult English language learners (ELLs) speaking fluency and accuracy. In this study, a series of RT scripts were developed and implemented with a group of low-intermediate level ESL learners. Pre and post-test quantitative and qualitative data were collected to determine the impact of RT on adult ELLs oral skills and self-confidence. A total of 12 ELLs ages 18 to 36 years old participated in an oral fluency class where a series of four different RT activities were implemented over the course of 14 weeks of instruction. Findings showed that RT activities not only enriched the L2 classroom experience, but learners were overwhelmingly positive about the impact of the technique in improving their general fluency, accuracy, and level of self-confidence in speaking English.

INTRODUCTION

Drama is an approach that has captured the attention of language teachers as a means to provide language learners with the opportunity to freely engage in interactive dialogue in a safe environment where they experience the language rather than merely learn about it. Drama can also provide students with the opportunity to practice real-life scenarios they may encounter helping a greater number of students become more confident and fluent in their communication (Burke and O'Sullivan, 2002, p. xix). Boudreault (2010) encourages the use of drama in language classes by stating, "ESL/EFL professionals need to use this medium more because the artificial world of the classroom can be transformed into a quasi-real language situation and provides an endless amount of opportunities for students' personal growth" (para. 1).

While there are many types of dramatic techniques that have been used in language classrooms such as skits, full-length plays, improvisation, etc., there is one particular technique that has the potential to capture learners' attention and develop measures of oral fluency while not requiring learners to move about a stage, use costumes, extensive props, or actions (Moran, 2006). This technique is Readers Theater (RT).

Background Information

Readers Theater (RT) was initially developed to help children develop their oral reading skills (Corcoran, 2005), including reading fluency (Millin and Rinehart, 1999). In RT, students take on the persona of a character in the script. They use facial expressions, hand gestures, and speak with great emotion to demonstrate their character's personality. Typically, these scripts have been based on fairy tales or folklore (Ng, 2008). While these types of texts can be fun and

interactive, Ng and Boucher-Yip (2010) argue that they can also be used to help readers strive for voice flexibility, good articulation, proper pronunciation, and projection. Research done by Young and Rasinski (2009) with monolingual elementary school students found that RT helped learners build their ability to make meaning in English as well as reading fluency and confidence. Further, they found that unmotivated students became motivated and struggling readers thrived (p. 11).

English language teachers have also used RT in developing English as a second language (ESL) and English as a foreign language (EFL) students' speaking fluency and accuracy (Liu, 2000; Tsou, 2011). Wu (2015) reported that RT had a profound influence in students' learning and retention of English idioms. Ng (2008) commented that RT scripts provide a rich source of comprehensible input in language that is natural and spoken. Evie Tindall (2012), describes RT as a "time-tested [approach]… [providing] English language learners with content tailored to their abilities in addition to [innumerable] opportunities to engage in meaningful interactions with language, content, and teachers and peers" (p. 36).

Even with the positive influences that research into RT has demonstrated, unfortunately, the majority of available scripts have been created for young learners and are based on stories largely appropriate for learners in grades K-6 (Garrett and O'Connor, 2010). The content of the scripts seldom relate to the themes, vocabulary, and textbook content that adult ESL learners are experiencing in their ESL classes. Besides the challenge of finding appropriate scripts for adult learners, teachers often avoid using dramatic techniques such as RT for several reasons: 1) because they lack training in drama or how to implement role play in the classroom, 2) because they fear they will not be able to find teacher-friendly material that explains how to put on a play or that provides actual scripts appropriate for classroom use, 3) that using dramatic techniques turns too much control over to the students and the class may get out of control (Burke and O'Sullivan, 2002, p. xxii) and finally, 4) because teachers often admit they lack training in how to organize instruction that will assist learners with developing elements of pronunciation and articulation in English (Foote, et al., 2011).

The focus of the current study was to use RT to help promote oral skills development in adult ESL learners enrolled in an intensive English program. A total of four RT scripts were created that aligned with the course content. Pre and post surveys were completed at the beginning and end of a 14-week semester allowing students to evaluate how the use of RT influenced their oral fluency, accuracy, and level of self-confidence. Each script was introduced and practiced for a period of three weeks after which the students' performance was videotaped and students viewed the tape for personal reflection and development, and each student received personal feedback from the teacher. The program's Speaking Level Achievement Test (LAT) was comprised of a series of 12 leveled speaking tasks. The test was administered at the beginning and end of the semester to both the treatment and control groups of students in an effort to determine whether the RT treatment enhanced the participants' overall speaking proficiency. An end-of-semester survey and focus groups were conducted in an effort to collect qualitative feedback from the treatment group about their experience with RT.

Research Questions

In an effort to empirically evaluate the influence of RT on the oral skills and self-confidence of intermediate-level ESL learners, the following research questions were addressed:

1. How did the ESL learners evaluate the impact of RT on their perceived levels of fluency, accuracy, and self-confidence?
2. What qualitative feedback did learners provide regarding the embedding of RT into the course curriculum?
3. Did the end of semester level achievement tests (LATs) show a difference in speaking skills between the control and treatment groups?

METHODS

To examine the influence that RT had on the ESL learners in a 14-week intensive English program, a quasi-experimental study was constructed. Participants in the control group did not receive any instruction involving RT. In the treatment group, however, learners participated in a 14-week instructional period, where every three weeks, students would practice a scripted dialogue involving multiple speakers. At the end of the three weeks of practice, the presentation was videotaped for the purpose of self-evaluation and teacher evaluation of the students' performance.

Participants

Subjects in the study were all studying English as a second language in a large intensive English program connected to a private university in the Western United States. At the beginning of the semester, students who scored at the intermediate level on the program's placement test were randomly placed into one of three Foundations C (intermediate-med) classes. Two classes (a total of 26 ELLs) were designated as the control group where they completed the normal curriculum. The third class (12 participants) was designated as the treatment group. They received 14 weeks of RT instruction in addition to the normal curriculum. A total of 61.5% of the ESL learners had spent one year or less studying English. The other students (38.5%) had studied English for more than a year. The age range for all of the participants was from 18 to 34 years old. For students in the treatment group, 84.6% of the students had Spanish as their first language (L1), and 7.7% of the students were Portuguese speakers and 7.7% were Chinese speakers.

Surveys

Treatment group participants completed pre and post surveys during the semester. The initial survey captured demographic data from the participants including their age, first language, gender, and length of time they had studied English. Near the end of the semester, students in the treatment group were asked to participate in a focus group. This focus group was facilitated by an experienced teacher who did not teach at the English language program. Students in the treatment group also signed consent forms at the beginning of the semester consistent with regulations put in place by the university's Institutional Review Board, that safeguarded students' identity and all data generated from the study.

Script Creation

Because a major purpose of this study was to empirically evaluate the use of RT embedded into an intermediate-level intensive English course, the RT scripts were created from topics and themes identified in the course textbook. The topics included: academic honesty, finding a job,

ways of travel, and modern innovations. There were a few guidelines that researchers used in creating the scripts. These guidelines included:

- The topic should be authentic and relevant to the learners' situation
- The script needed to include idiomatic expressions and key vocabulary from the unit lessons
- The scripted dialog needed to include a conversation between four to five different characters of different genders, personality types, and ages
- The scripted dialog was to contain a conflict or problem that speakers in the conversation would seek to resolve in the course of the dialog
- The dialog needed to include language that was emotive, requiring the use of facial expressions, hand gestures, and inflections in the voice
- The script length should allow for a minimum of six to eight speaking turns per character which meant that the dialogs were 35 to 40 lines long

Once the scripts were created, each was piloted with a small group of native English speakers to ensure a natural flow in the language, the use of a variety of emotions, and appropriate phrasing and wording consistent with how people would engage in authentic dialog. After appropriate adjustments were made, the group of native speakers then rehearsed the dialog so that they could perform it naturally including times when speakers might be speaking at the same time as another person which often happens in real speech. After several minutes of rehearsing, the dialog was then recorded for use by the teacher and ESL learners inside and outside the classroom. This same process was used in recording each of the four scripted dialogs.

Implementation of the Scripts

It is important to note that during the first week of the course and prior to the implementation of the first script, the teacher reviewed and practiced with students several basic features of effective pronunciation. These features included the importance of clear articulation, the purpose of pausing when producing chunked expressions, the role that word and sentence stress have in emphasizing important syllables and words in a sentence, and the use of rising and falling pitch when a speaker produces different types of utterances such as yes/no questions, wh-questions, statements, and expressions that contain a high degree of emotion. Once students were aware of the importance and use of prosodic features, the teacher could utilize this knowledge in implementing the scripts.

The cycle that was followed by the teacher in implementing each of the four scripts into the class is given in Figure 1.

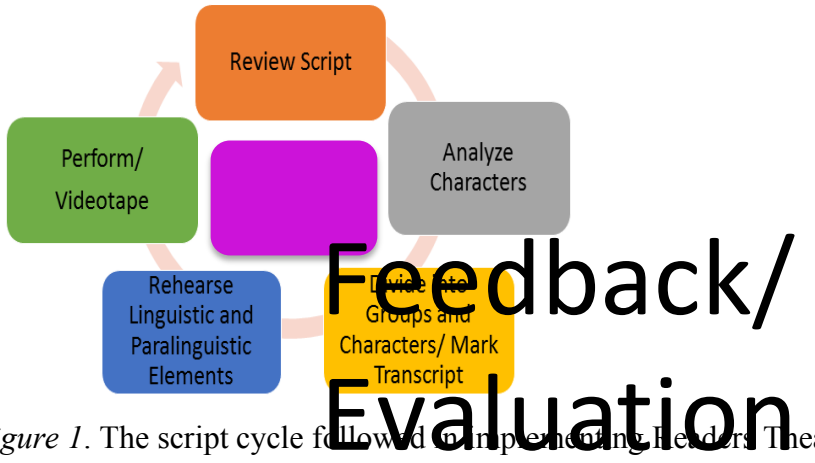


Figure 1. The script cycle followed in implementing Readers Theater into the class.

Each cycle lasted a total of three weeks and provided the students the opportunity to move through several steps including: 1) reviewing the script for new vocabulary, idiomatic expressions, and grammar; 2) analyzing the characters including their attitudes and emotions; 3) marking the script for prosodic features such as pausing, word and sentence stress, and intonation contours; 4) practicing the dialog both in class and at home with the markings and the native English speaker recording to develop a smooth flow in the delivery of each line; and 5) having the final performance videotaped for self-assessment and teacher evaluation.

As students practiced the script, the teacher provided regular feedback throughout the process by assisting with prosodic accuracy, the use of timing and turn-taking, facial expressions and hand gestures, and incorporating sufficient volume, energy, and emotion in the discourse.

RESULTS

Quantitative Findings

Participants in the treatment group were asked in the pre and post surveys, to identify on a 5-point Likert scale their perceived levels of fluency, accuracy, and self-confidence at that respective point in time. The feedback provided showed that in all three areas, fluency, accuracy, and self-confidence, students felt they improved dramatically over the course of 14 weeks (See Figures 2, 3, and 4).

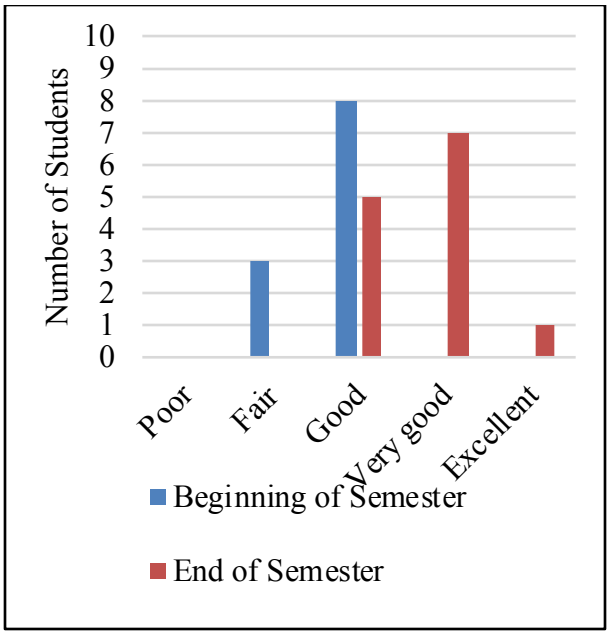


Figure 2. Change in ESL students' perceived level of fluency (pre & post)

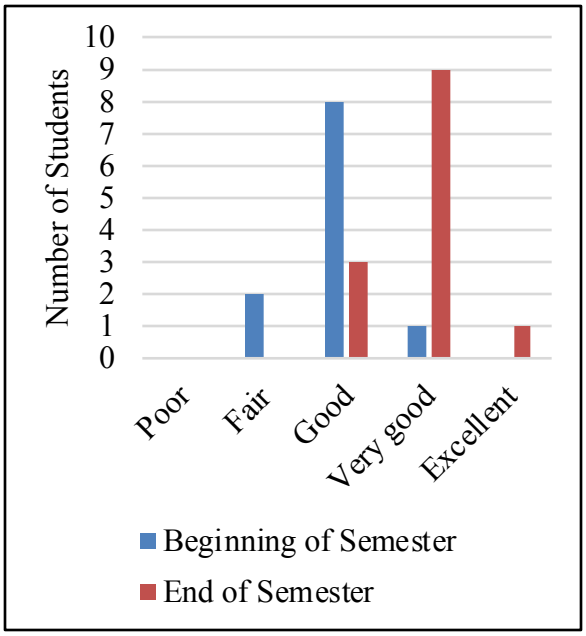


Figure 3. Change in ESL students' perceived level of accuracy (pre & post)

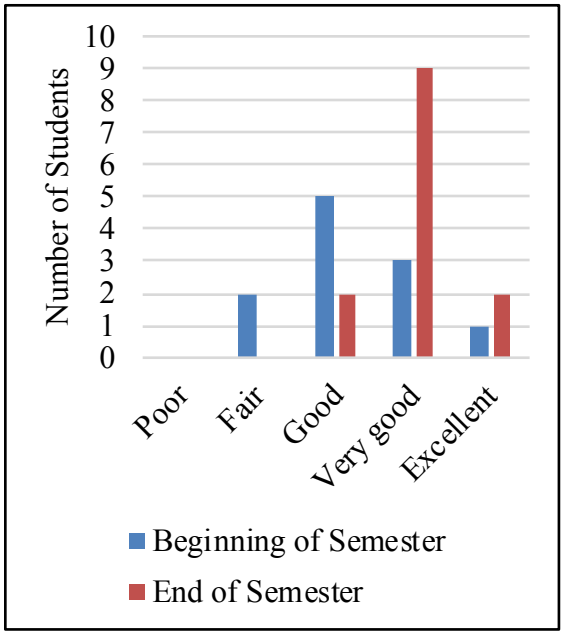


Figure 4. Change in ESL students' perceived level of self-confidence (pre & post)

The students were also asked to evaluate specifically how RT helped improve their speaking, listening, and pronunciation skills on a scale from 0 to 10, with zero being "not at all" and 10 being "extremely improved." When calculating the mean score for all students, the influence of RT on speaking skills was 9.5, listening was rated at 6.25, and pronunciation was at a level 10. Finally, the level achievement test scores (LATs) of students in the control and treatment groups were compared from the beginning of the semester to their test scores at the end of the semester.

In the pre-test, the average scores for the control and treatment groups only differed by a measure of 0.02. The control group's performance was at 3.26 on a proficiency scale ranging from 0 to 7. The treatment group's score was at 3.24. At the end of the semester, both groups improved significantly over time as determined by an analysis of variance (ANOVA) ($F(1,37) = 12.644$, $p = .001$. The partial eta squared =.26 (from the pre-test to the post-test) showed a large effect size.) However, when comparing the between group scores from the pre-test to the post-test, there was no significant difference between the groups ($F(1,37) = 1.365$, p=.250, partial eta squared = .037).

Qualitative Findings – Focus Groups

In addition to the Likert scale data from the pre- and post surveys, two focus groups were held outside of class time in which a total of 11 students from the treatment group participated. In probing the participants more specifically regarding the use of RT in the classroom, they stated on a scale from 0 to 10 with 10 meaning the activity was extremely beneficial to their English development, RT was rated at a value of 8.5. With regards to the value of RT in the intensive ESL classroom, all focus group participants stated that they felt the activity was fun and the teacher helped make it so. The participants also commented that RT provided focused practice on pronunciation, fluency, stress, pausing, and intonation. They also appreciated learning the idiomatic expressions that were built into the conversations. Participants stated that they felt the technique had helped build their self-confidence in speaking English and they enjoyed listening to and watching their videotaped conversations on YouTube.

The focus group participants also provided some helpful recommendations for future application of RT in the ESL classroom. Recommendations included: 1) Practicing additional scripts dealing with situations occurring in every day contexts like communicating at the bank, talking with doctors and nurses at the hospital, or shopping; 2) Making the scripts longer with additional turns; 3) Progressively shortening the time it took to prepare the scripts from three weeks to two weeks or even one week; 4) Memorizing the final script and performing it for another class. In this way, students suggested that the final performance would be more of a real performance since it would be for others and not just classmates.

DISCUSSION

As previous research (Jordan and Harrell, 2000; Ng, 2008; Liu 2000) has shown, RT can be a fun, interactive activity that involves ESL learners in oral discourse practice both inside and outside the classroom. It provided learners with the opportunity to be engaged in focused practice on the use of prosodic features in developing oral fluency and accuracy through the conversations rehearsed in the RT scripts. Learners also reported that the RT activities developed their self-confidence in speaking English, and facilitated their connection of linguistic and paralinguistic features through the scripted dialogs. The videotaped performance further provided students and teacher a means of evaluating individual student performance in the areas of speech rate, emotive language, appropriate phrase and sentence stress, intonation, and chunking of language through the appropriate use of pausing and thought groups.

CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH

Even though there was a small number of students involved in the treatment group and the treatment only lasted for 10-15 minutes a day over the course of 14 weeks, participants did report dramatic improvement in areas of oral fluency, accuracy, and self-confidence due to the use of RT in the classroom. To better determine the actual influence of RT on ESL students' linguistic change in English, it may be necessary to design specific read aloud tasks in the LATs in addition to the spontaneous speech tasks to accurately measure specific changes in students' fluency and accuracy. These tasks could be designed so that students' speech rate, articulation rate, pause frequency, stress placement, and use of T-units could be measured. Currently, the LATs are rated holistically using a speaking rubric developed in-house for student placement. A pronunciation rubric (Ma, et al., 2018) could be used as well to better assess learners' pronunciation development as it may be impacted more directly through the RT tasks.

The positive results from this study suggest that additional inquiry into the use of RT with adult ESL learners needs to be further explored. The positive reactions of learners to this technique suggests that at a minimum, the activity provides a fun, interactive, and yet challenging means for learners to improve linguistic and paralinguistic features present in conversational discourse, something that all students need to be able to perform as competent speakers of a second language.

ABOUT THE AUTHORS

Mark Tanner is an assistant professor in the Linguistics Department at Brigham Young University. His research focuses on second language pronunciation and pedagogy, fluency and comprehensibility, and teacher training. He has taught ESL and EFL for 40 years and teaches undergraduate and graduate TESOL courses at BYU.

Alisha Chugg is an ESL instructor, teacher trainer, and coordinator at Internexus, a language school in Provo, Utah. She has taught ESL for 7 years and is particularly interested in helping her students achieve their goal of fluency and accuracy in English.

Contact Information
Mark Tanner
Brigham Young University
4063 JFSB, Provo, UT 84602
(801) 422-8154
Email: Mark_Tanner@byu.edu

Alisha Chugg
Email: alchugg@gmail.com

REFERENCES

Burke, A. & O'Sullivan, J. (2002). *Stage by stage: A handbook for using drama in the second language classroom.* Portsmouth, NH: Heinemann.

Corcoran, C. (2005). A study of the effects of readers' theater on second and third grade special education students' fluency growth. *Reading Improvement, 42*, 105-111.

Foote, J., Holtby, A., & Derwing, T. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29*, 1-22.

Garrett, T. & O'Connor, D. (Sept/Oct. 2010). Readers' theater: "Hold on, let's read it again." *Teaching Exceptional Children, 43*, 1.

Jordan, S. & Harrell, K. (2000). Readers Theatre: A creative tool for strengthening skills of emergent readers. *Kindergarten Education: Theory, Research and Practice, 5*, 73-80.

Liu, J. (2000). The power of RT: From reading to writing. *ELT Journal, 54*, 354-361.

Ma, J., Henrichsen, L., Cox, T., and Tanner, M. (2018). Pronunciation's role in English proficiency ratings. *Journal of Second Language Pronunciation, 4,* 73-102.

Millin, S. & Rinehart, S. (1999). Some of the benefits of RT participation for second-grade Title I students. *Reading Research and Instruction, 39*, 71-88.

Moran, K. (2006). Nurturing emergent readers through RT. *Early Childhood Education Journal, 33*, 317-323.

Ng, P. (2008). *The impact of RT (RT) in the EFL classroom.*[PDF document]. *Polyglossia, 14*, 93-100. Retrieved from http://www.apu.ac.jp/rcaps/uploads/fckeditor/publications/polyglossia/Polyglossia_V14_Ng.pdf

Ng, P. & Boucher-Yip, E. (December, 2010). RT: Improving oral proficiency in a Japanese university EFL course. ELTWorldOnline.com. Retrieved from http://blog.nus.edu.sg/eltwo/2010/12/20/readers-theatre-improving-oral-proficiency-in-a-japanese-university-efl-course/

Tsou, W. (2011). The application of RT to FLES (Foreign language in the elementary schools) reading and writing. *Foreign Language Annals, 44*, 727-748.

Wu, S. (March, 2008). Effective activities for teaching English idioms to EFL learners. *The Internet TESL Journal, 14*(3). Retrieved from http://iteslj.org/Techniques/Wu-TeachingIdioms.html

Young, C. & Rasinski, T. (2009). Implementing Readers Theatre as an approach to classroom fluency instruction. *The Reading Teacher, 63*(1), 4-13.

Violin-Wigent, A. (2018). The impact of explicit instruction on the pronunciation of French liaisons. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 194-203). Ames, IA: Iowa State University.

THE IMPACT OF EXPLICIT INSTRUCTION ON THE PRONUNCIATION OF FRENCH LIAISONS

Anne Violin-Wigent. Michigan State University

In order to determine if the explicit description of the many rules traditionally given to explain French liaisons have an impact on students' production of liaisons, I compare seven recordings made by 25 undergraduate students enrolled in an intact third-year class on French pronunciation. Recordings were coded for the accurate production of obligatory and prohibited liaisons and results were analyzed using Goldvarb. The analysis in this pilot study shows that liaisons are pronounced correctly around 80% of the time. Explicit instruction seems to have an effect since the first recording (before instruction) is associated with the lowest rate of accurate production while the sixth recording of the semester is associated with the highest. The trajectory of the improvement, however, is not linear with several setbacks during the semester. Another statistically significant factor is the syntactic environment of the liaison. Some environments show a ceiling effect with high accuracy from the beginning. Others shows gains over the course of the semester while some don't. The discussion centers on which of these contexts seem to improve most after explicit instruction and should, therefore, be included in overt explanations of liaisons.

INTRODUCTION

In order to measure the effectiveness of explicit instruction, this pilot study focuses on the accurate realization of obligatory and prohibited liaisons. Liaison is a French sandhi phenomenon that can be defined as follows: in some contexts, the otherwise-silent final consonant of a word will become pronounced as the onset of the following word when this following word starts with a vowel. For example, we note a contrast in the two pairs below, where the vowel-initial word in the 'friend' example starts with a consonant that is not normally pronounced at the end of the determiner (as shown in the left-hand examples below), and that is resyllabified as the onset of the word for 'friend'.

(1)	les filles ('the girls')	les amies ('the friends' feminine)
	[**le**.fij]	[le. ‿**Z**a.mi]
(2)	un garcon ('a boy')	un ami ('a friend' masculine)
	[**ɛ̃**.gaʁ.sõ]	[ɛ̃. ‿**N**a.mi]

Liaisons are described as suprasegmental by Kennedy, Blanchet, and Trofimovich (2014) and Gordon and Darcy (2016) since they involve linking between words. But at the same time, since they involve the presence or absence of a particular consonant, they can also be described as a segmental phenomenon.

According to traditional descriptions, starting with Delattre (1951) and many since then, there are 3 'types' of liaison, classified according to their frequencies: obligatory liaisons are said to occur 100% of the time when the conditions are met, prohibited liaisons never occur, and

optional ones can be pronounced or not depending on non-linguistic factors such as style. All three types of liaisons are typically taught very explicitly with a list of all the contexts, yielding five different rules for obligatory liaisons, nine for prohibited, and six for optional liaisons in Violin-Wigent, Miller, and Grim (2013), the textbook used for the class analyzed in this study. Because of this, they provide a good testing ground for an investigation on the efficacy of explicit pronunciation instruction. After a brief review of previous studies on this issue, the methodology adopted in this study will be described before results are presented and discussed.

EXPLICIT PRONUNCIATION INSTRUCTION

Many studies have investigated the question of explicit instruction (EI) in the acquisition of pronunciation, such as Kissling (2013 and 2014), Saito (2013 and 2015) and Gordon and Darcy (2016), among many others. In an early study, Derwing and Munro (2005: 387-8) conclude that EI is beneficial because it helps students notice the differences between the target and their own production. In their meta-analysis of 86 studies, Lee, Jang, and Plonsky (2015) show a significant effect of EI but they nuance their conclusion by suggesting that studies without significant results are probably not published. Thomson and Derwing (2015: 332-3) reviewed 75 studies and also noted the positive results of EI in 82% of cases when learners read texts aloud, rather than produced spontaneous speech.

Most published studies center on English (as a foreign or second) languages. Methods and results are not always transferable to other languages, as a fair number of them focus on suprasegmental phenomena (particularly on stress placement), which may not apply to other languages. Among studies on other languages, we can cite, among others, Chun, Jiang, and Avila (2013) for Chinese, and, for Spanish, Elliot (1995), Gonzales-Bueno (1997), Lord (2005, who describes improvement during a Spanish phonetics class, but without a control group), Gonzales-Bueno and Quintana-Lara (2011) and Bajuniemi (2013). As far as French is concerned, we can cite Weinberg and Knoerr (2003) whose primary interest is in including technology, and Ruellot (2011) on the distinction between [u] and [y]. French liaisons are studied in Kennedy et al. (2014), who show a relationship between instruction, awareness, and production in an FSL context. They classify liaisons under quantitative awareness, which is associated with memorization of rules or chunks. In addition, De Moras' studies (2011 and 2013) report on the effect of three types of treatment (instruction, repetition, and feedback) given to learners in different orders. She concludes that explanations are not very beneficial in the acquisition of liaisons and that repetitions seems much more valuable.

METHODOLOGY

In this context, the investigation focuses on the evolution of learners in a class focused on French pronunciation and phonetics. The class, taught at a large Midwestern university, includes 25 students and can be considered an intact class since the productions at the basis of this study were part of their regular class assignments. These "oral quizzes" are seven recordings that the learners enrolled in the class turned in during the 15-week semester. The lesson on liaisons is associated with recording 4, considered as the immediate post-treatment test.

Students read texts aloud, recorded them using Audacity, and uploaded them on the class website. The task of reading aloud was chosen because it reflects what was done in class and this

study focuses on the effectiveness of in-class instruction. As described in the washback effect (Lee & VanPatten 2003, p. 100), it is necessary that testing reflects classroom activities, and vice-versa. We also follow the assumption that if liaisons are not produced accurately in a guided activity with heavy self-monitoring and a focus on accuracy, they will not be produced in more spontaneous contexts. Since this is the first class where liaisons are discussed, we also assume that they will not have been acquired yet. In addition, as mentioned by Tomson and Derwing (2015, p. 339), spontaneous tasks should be reserved for longitudinal studies as explicit teaching done in the phonetics class is more akin to awareness raising while true learning or acquisition will take place beyond the time-limit of the class. Finally, this was a way to eliminate learner avoidance strategies. For example, in the case of common liaisons, like between articles and nouns or pronouns and verbs, it is not hard to select a different noun or verb that starts with a consonant to avoid a liaison, or simply to pause between the two elements, which would automatically block the liaison. The task of reading aloud, however, could be detrimental to accurate production since a task based on the graphical representation of words may reinforce or at least trigger two types of mistakes that are found to be common: the non-linking of the liaison consonant to the following word (*non-enchaînement*) and the spelling pronunciation of the liaison consonant (such as [d] instead of [t] in words such as *quand* 'when').

Following Liakin, Cardoso, and Liakina (2017: 356), two advanced learners of French served as research assistants. After training and discussion, they listened to each student's recordings for accuracy of liaison production. In the case of conflicting transcriptions, the author made the final decision. The inter-rater reliability rate was around 98%. Liaisons produced accurately compared to a native-like target (following the nativeness principle) were coded with 1. All others were coded as 0, including liaisons that were not produced when they should have been or produced when they shouldn't have been, liaisons that were produced with the wrong consonant, and liaisons in which the liaison consonant was not resyllabified.

The recordings produced a total of 1625 potential sites for obligatory and prohibited liaisons. However, 131 tokens were excluded for one of the following reasons: students didn't turn in a recording; words or sentences were skipped or transformed/misread; or students made an optional pause between words in cases of prohibited liaisons. Statistical analysis was carried out using Goldvarb, a logistic regression program that assigns a weight to each element to describe its influence on the probability of choosing each option in a pair (in this case, the accurate vs. non-accurate production of liaisons). In the tables below, higher weights indicate that the associated element promotes or increases the likelihood of accurate liaisons while lower weights indicate that the associated element demotes or decreases the likelihood of accurate liaisons. While unusual in the field of language acquisition, Goldvarb is widely used in sociolinguistics as it is designed to compensate for very uneven numbers of tokens in cells, which is the case here. Four factors were tested: time in the semester (before treatment, immediately after treatment, after treatment), the recording number (from 1 to 7), the type of liaison (obligatory vs. prohibited), and the context of liaison (corresponding to the explicit rules given to the learners in class).

RESULTS AND ANALYSIS

Overall results show that 80.4% of the remaining 1494 liaisons were produced accurately. Among the four factors tested, only two were selected as significant by Goldvarb: the recording number and the type of liaison.

Recording Number

Table 1 below presents the results for the accurate pronunciation of liaison according to the recordings done by learners during the semester.

Table 1

Results for recording number

	Relative weight	N correct	N total	Percentage
1	0.316	122	195	62.6
2	0.467	187	220	85
3	0.196	54	99	54.5
4 (treatment)	0.443	210	273	76.9
5	0.654	121	133	91
6	0.644	267	291	91.8
7	0.610	240	283	84.8
	Range = 0.458			

Before discussing the results in this table in detail, we should mention that the weight associated with recording 3 may not be representative of learners' actual liaison performance of the students for several reasons. First, this recording only contained four possible tokens of liaisons, which is a fairly low number to be able to generalize from. In addition, two of these tokens involved the phrase *cent autres (renards)* '100 other (foxes)'. As will be explained in section 4.3 below, liaisons after numbers are particularly problematic when the consonant involved is not a [z]. Indeed, native speakers consulted on this mentioned that they would prefer to rephrase it to avoid a liaison that was deemed odd. Ignoring the figure for recording 3 then, the relative weights presented in Table 1 show a steady increase between recording 1 and recording 5, with a small set back immediately following treatment (on recording 4), and a more important slide back after the peak in recording 5. This indicates that, in spite of a small delay, some improvement has taken place after treatment. Indeed, the weights associated with recordings 5, 6, and 7 are all above 0.5 (which is considered neutral), hence showing that these recordings promote the production of accurate liaisons. At the same time, the weights associated with the recordings before recording 5 show weights below 0.5, therefore indicating that liaisons in these recordings were more likely to be produced incorrectly.

While these results are encouraging, we notice that the range of weights (in the last row) is relative small, especially when compared with the range of the weights associated with the contexts of liaisons (in Table 2 below). This indicates that, while significant, the recording

number is not the most important factor in explaining the variation in the accurate pronunciation of liaisons, but that the context of liaison is.

Context of Liaison

Table 2 below presents the results for the accurate pronunciation of liaison according to the contexts in which liaisons occur. These do not reflect all possible contexts but only the ones present in the recordings.

Table 2

Results for context of liaison

	Relative weight	N correct	N total	Percentage
Before or after 'et'	0.758	213	224	95.1
In fixed phrases	0.730	24	25	96
Between two groups	0.732	214	226	94.7
Between a pronoun and a verb	0.468	97	120	80.8
Between a determiner and a noun	0.460	408	534	76.4
Between a subject NP and a verb	0.378	114	137	83.2
After a monosyllabic adverb	0.322	31	50	62
After a monosyllabic preposition	0.244	68	93	73.1
After a monosyllabic conjunction	0.058	17	70	24.3
	Range = 0.700			

As can be seen in the first few rows in this table, three contexts seem to be acquired fairly well since their weights are very high, reflecting a high likelihood of accurate production. These contexts include two prohibited liaisons (before or after the conjunction *et* 'and' and between two groups) as well as one obligatory, in the fixed expression *peut-être* ('maybe'). In addition, it should be noted that all the cases (15 tokens) involving a prohibited liaison before a proper noun were realized correctly and are, therefore, not shown in the chart. Additional statistics were run on the two prohibited contexts to see if there was a change over the duration of the semester. These gave non-significant results, showing that these two contexts of liaison were acquired even before the class began. These results lend support to Kennedy et al.'s (2014) classification of liaisons under quantitative awareness where contexts that are more chunk-like, like fixed phrases or around a particular word, seem to be acquired first.

Table 2 also shows two contexts that are fairly neutral, namely, between a clitic pronoun and a verb and between a determiner and a noun. Finally, all the other contexts may be deemed

problematic since their weights indicate that they are very likely to be associated with inaccurate production of liaisons. Most of these contexts involve an obligatory liaison after a monosyllabic function word. A closer look at these categories is needed to investigate whether improvement or change has taken place as a result of instruction.

Evolution of Non-Acquired Contexts

Starting with the two contexts associated with neutral weights in Table 2, no significant effect of treatment or recording was found between a clitic pronoun and a verb, hence showing no evolution during the semester. A tendency can be seen however, as percentages show a rise from 72.3% to 86.3% before and after the lesson. Table 3 below present the evolution of the results for determiners depending on the lesson.

Table 3

Evolution of liaisons after determiners according to treatment

	Relative weight	N correct	N total	Percentage
Pre-lesson	0.382	173	266	65
Lesson	0.434	104	125	83.2
Post-lesson	0.756	131	143	91.6
	Range = 0.374			

This table suggests that instruction had a positive effect on the correct production of the liaisons after a determiner, even though this effect was not immediate as can be seen in the weight associated with the lesson, which is below still 0.5 but higher than before the lesson. By the end of the semester, liaisons after a determiner seem almost fully acquired as the percentage approximate those from Table 2 that were deemed acquired. Two additional factors were tested regarding determiners (the type of determiner and the liaison consonant) but the former was found to be significant, as shown in table 4 below.

Table 4

Results according to the type of determiner

	Relative weight	N correct	N total	Percentage
Indefinite plural *des*	0.780	90	96	93.8
Definite plural *les*	0.670	111	122	91
Possessive singular *mon*	0.525	121	147	82.3
Indefinite *quelques*	0.436	16	23	69.6
Indefinite singular *un*	0.244	36	71	50.7
Numerals *six* and *cent*	0.141	34	75	45.3
	Range = 0.639			

Table 4 clearly shows a preference for liaisons with plural determiners, with the exception of *quelques* (‘a few’), which is an unusual determiner in French in so far as it is multisyllabic. This is not surprising as liaisons with [z] are very strongly associated with the plural in French, so strongly in fact that false-liaisons are known to occur after numbers that do not end with an orthographic <s> such as *quatre* (‘four’) or *huit* (‘eight’), or even *cent* (‘one hundred’). This may help explain the low weight associated with the last category since two of the three tokens for each student involve *cent* (realized accurately 28% of the time vs. 80% for *six*). Additional data is need in this category to come to a more definite conclusion.

Table 5 below presents the results for the only significant factor retained for the context of between a subject NP and verb. It is to be noted that this context was not included in the recordings prior to the lesson and that only one token per student was in the recording associated with the lesson. For this reason, as well as the fact that the percentages and the weights do not show the same progression, we need to be prudent in our analysis.

Table 5

Evolution of subject NP and verb according to recording

	Relative weight	N correct	N total	Percentage
4 (treatment)	0.118	12	25	48
7	0.685	42	48	87.5
6	0.503	60	64	93.8
	Range = 0.567			

This table shows an improvement in the realization of liaison but not immediately after the lesson. It would have been beneficial to have not only tokens prior to the lesson but also for recording 5 in order to have a more complete picture of the evolution of students in this context.

As far as the last three contexts in Table 2 are concerned (after monosyllabic adverbs, prepositions, and conjunctions), we need to be cautious with the results due to small numbers. In particular, there were only two tokens per student of an adverb, one in recording 1 with 36% of correct pronunciation of liaisons and one in recording 4 (associated with the lesson) with 88% accuracy. Four prepositions per student are included in the corpus, one in recording 2 (produced accurately 83.3% of the time), one in recording 5 (87.5% accuracy) and two in recording 6 (60% accuracy). Finally, three conjunctions are found in the following three recordings: one in recording 1 (12% accuracy), one in recording 4 (17.4% accuracy), and one in recording 6 (45.5% accuracy). These figures seem to indicate that there is an improvement as the semester progresses, except for the back sliding observed with prepositions in recording 6. Again, tokens are too few and too unevenly spread out during the semester to come to a definite conclusion for these three contexts.

CONCLUSION

Even without comparing these results with those from a control group, we can conclude that there is a significant improvement of accuracy over the course of the semester. The trajectory of

the improvement, however, is not linear as the second recording (still before instruction) is associated with a higher accuracy than the recording linked with the liaison unit, hence suggesting that students have improved on their own, but regressed with the introduction of the explicit rules governing liaisons, showing a pattern of back sliding, especially with the final recording. We have also observed variation in accuracy depending on the context of liaisons. Indeed, three contexts show a ceiling effect: around *et*, in idiomatic expressions, and between two groups, while almost all the other contexts show a positive evolution, with the exception of between a clitic pronoun and a verb.

These results seem to contradict those from De Moras (2011 and 2013) but several elements need to be mentioned to qualify this statement. First of all, students in her studies all received treatment but in different orders and her conclusions are really about the optimal order of presentation rather than the effectiveness of explicit instruction. In addition, neither study has a control group that would enable us to see the effect of treatment as opposed to the effect of additional language exposure without any focus (feedback, repetitions, or explicit instruction) on liaison. I am currently collecting data from such a control group for a future study.

Finally, though preliminary, the current results suggest that class time should include more attention to contexts where improvement is possible (between determiners and nouns) and on difficult contexts (such as between clitic pronouns and verbs and after prepositions, conjunctions, and adverbs). Needless to say, further research is needed on these contexts as well as on others not included here or without enough tokens to bring valid results. In addition, further research should compare tasks to investigate if results from reading-aloud tasks transfer to spontaneous speech, including in longitudinal studies.

ACKNOWLEDGEMENTS

Many thanks to the MSU College of Arts and Letters Faculty Research Award for funding my RAs.

ABOUT THE AUTHOR

Anne Violin-Wigent is an Associate Professor of French at Michigan State University. Her research focuses on French phonetics, in particular, on the effectiveness of explicit instruction in the acquisition of L2 pronunciation, and on the leveling of regional accents in France. She co-authored a textbook on French pronunciation *Sons et Sens: La prononciation du français en contexte* (2013) with Jessica Miller and Frédérique Grim.

Michigan State University
619 Red Cedar Road, Wells Hall B331
East Lansing, MI 48824
517- 884-6304
violinwi@msu.edu

REFERENCES

Bajuniemi, A. (2013). Teaching intervention on the pronunciation of Spanish intervocalic /d/. In C. Howe, S. E. Blackwell, & M. L. Quesada (Eds.), *Selected Proceedings of the 15th Hispanic Linguistics Symposium* (pp. 39-50). Sommerville, MA: Cascadilla Press.

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

Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains*. Middlebury College, VT.

De Moras, N. (2011). *Acquisition de la liaison et de l'enchaînement en français L2: Le rôle de la fréquence*. Electronic Thesis and Dissertation Repository, Paper 176. University of Western Ontario.

De Moras, N. (2013). The role of pronunciation instruction on the acquisition of *liaisons* by Anglophone speakers. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference.* (pp. 167-181). Ames, IA: Iowa State University.

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39*(3), 379-397. doi: 10.2307/3588486

Elliot, A. R. (1995). Foreign language phonology: Field independence, attitude, and the success of formal instruction in Spanish pronunciation. *The Modern Language Journal, 79*(4), 530-542.

Gonzales-Bueno, M. (1997). The effects of formal instruction on the acquisition of Spanish stop consonants, *Contemporary Perspectives on the Acquisition of Spanish, 2*, 57-75.

Gonzales-Bueno, M. & Quintana-Lara, M. (2011). The teaching of second language pronunciation through processing instruction. *Applied Language Learning, 21*(1-2), 53-78.

Gordon, J., & Darcy, I. (2016). The development of comprehensible speech in L2 learners: A classroom study on the effects of short-term pronunciation instruction. *Journal of Second Language Pronunciation, 2*(1): 56-92 doi: 10.1075/jslp.2.1.03gor

Kennedy, S., Blanchet, J., & Trofimovich, P. (2014). Learner Pronunciation, Awareness, and Instruction in French as a Second Language. *Foreign Language Annals, 47*(1), 79-96. doi: 10.1111/flan.12066

Kissling, E. M. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *The Modern Language Journal, 97*(3), 720-744. doi: 10.1111/j.1540-4781.2013.12029.x

Kissling, E. M. (2014). What predicts the effectiveness of foreign-language pronunciation instruction? Investigating the role of perception and other individual differences. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 70*(4), 532-558. doi:10.3138/cmlr.2161

Lee, J., & VanPatten, B. (2003). *Making communicative language teaching happen*. Boston: McGraw-Hill.

Lee, J., Jang, J., & Plonsky, L. (2015) The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics,* 36(3), 345–366. doi: 10.1093/applin/amu040

Liakin, D., Cardoso, W., & Liakina, N. (2017) The pedagogical use of mobile speech synthesis (TTS): focus on French liaison. *Computer Assisted Language Learning, 30*(3-4), 348-365. doi: 10.1080/09588221.2017.1312463

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania, 88*, 557–567. doi: 10.2307/20063159

Ruellot, V. (2011). Computer-assisted pronunciation learning of French /u/ and /y/ at the intermediate level. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*. (pp. 199-213). Ames, IA: Iowa State University.

Saito, K. (2013). Reexamining effects of form-focused instruction on L2 pronunciation development: The role of explicit phonetic information. *Studies in Second Language Acquisition 35*(1), 1-29. doi:10.1017/S0272263112000666

Saito, K. (2015). Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /r/ without explicit instruction. *Applied Psycholinguistics 36*(2), 377-409. doi:10.1017/S0142716413000271

Thomson, R.I., & Derwing, T.M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics, 36*(3), 326–344. doi: 10.1093/applin/amu076

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

Weinberg, A., & Knoerr, H. (2003) Learning French pronunciation: Audiocassettes or multimedia? *CALICO Journal, 20*(2), 315-336.

Wisniewska, N. & Mora, J. C. (2018). Pronunciation learning through captioned videos. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 204-215). Ames, IA: Iowa State University.

PRONUNCIATION LEARNING TROUGH CAPTIONED VIDEOS

Natalia Wisniewska, Universitat de Barcelona

Joan C. Mora, University of Barcelona

The current study investigates L2 learners' skills at integrating auditory and orthographic input while reading dynamic texts in L2-captioned video, as part of a broader research project investigating the role of exposure to L2-captioned video in L2 pronunciation development. Within this broader research goal, the eye movements of L1-Catalan/Spanish learners of L2-English (*n*=38) were recorded while watching short L2-captioned video clips. The Reading Index for Dynamic Text (Kruger & Steyn, 2013) was used as a measure of learners' amount of text processing, and an index of text-sound integration was computed by calculating the extent to which fixations on selected words synchronized with their auditory onsets. We also explored learners' individual differences in text-sound integration through a novel task that required learners to uncover text-sound mismatches. In addition, we measured learners' L2 segmentations skills through a word-spotting task (McQueen, 1996) and L2 proficiency through an Elicited Imitation Task (Ortega et al., 2002). The results shed light on the relationship between reading and audio-text integration skills, suggesting that efficient reading might be what leads to modality integration.

INTRODUCTION

Unlike other aspects of foreign language teaching (e.g. grammar, vocabulary), pronunciation does not often receive enough attention in second language (L2) classroom settings. It has been argued that pronunciation instruction had become a "casualty of Communicative Language Teaching" (Thomson & Derwing, 2015: 326). The problem with the pedagogical implementation of pronunciation instruction is twofold. First, it is not easy to embed a focus on pronunciation within a communicative teaching approach that emphasizes interaction (Mora & Levkina, 2017). Second, when pronunciation training takes place, it typically does so at an individual level, provided the learner has an interest and the willingness to seek out-of-the-classroom sources of input. One common source of exposure to L2 spoken input is watching movies in the L2.

Research has shown the pedagogical benefits of multimedia learning and the use of audio-visual materials (Mayer, 2009), such as subtitled video (Danan, 2004). Studies investigating the pedagogical potential of multimodality report improvement in listening comprehension (e.g., Vanderplank, 1988) and L2 vocabulary acquisition (e.g., Montero Pérez, Van Den Noortgate & Desmet, 2013). Such gains are explained by the notion of bimodal reinforcement in Paivio's (1986) Dual Coding Theory, which claims that the dual processing of auditory and visual information helps create and strengthen the mental representations of perceived objects, and therefore, promote learning. Could subtitled videos also enhance L2 pronunciation development?

Speakers of English as a foreign language at all levels of proficiency experience difficulty understanding words in a continuous stream of speech. For example, Charles and Trenkic (2015) asked L2 speakers in long-term-immersion to perform a shadowing task (Mitterer & McQueen, 2009) where they were asked to repeat back audio excerpts from TV programs. They found that L2 speakers failed to repeat about 30% of what they heard. This may be due to the challenging task of making use of L2-specific word boundary cues to identify words

(McQueen & Cutler, 1998) or to a likely mismatch between the incoming auditory input and the inaccurate phonological representations of L2 words in the learners' mental lexicon (Broersma, 2012). It is therefore important to investigate what sources of input exposure may have a positive effect on learners' improvement in L2 speech processing and comprehension.

Bimodal input in the form of subtitles can benefit pronunciation development in at least two ways. At the perceptual level, subtitles can aid the decoding and segmentation of speech by helping listeners map auditory input to linguistic form in running speech. The simultaneous presentation of word forms in written and auditory modalities facilitates word identification and lexical access and can boost speech processing (Mitterer & McQueen, 2009). We propose that efficiency in matching auditory and orthographic representations of words during extensive viewing would not only lead to gains in L2 speech processing, but it might also trigger changes in the phonological representation of words in the mental lexicon, eventually leading to more target-like pronunciation. Indirect evidence of the potential of bimodal input exposure for L2 phonological development is provided by Mitterer and McQueen (2009). In their study, Dutch participants exposed to English subtitles while watching videos in unfamiliar Scottish or Australian accents, outperformed participants who watched the videos without subtitles or with subtitles in the L1 on a shadowing task measuring L2 speech segmentation. They found that L2 subtitles enhanced foreign speech perception by boosting lexically-guided learning. Thus, extensive exposure to bimodal input provided by L2-captioned video may facilitate the development of more accurate L2 phono-lexical representations, eventually leading to improvement in L2 pronunciation.

The current study is a part of a larger project on the development of L2 pronunciation through L2 captioned videos, where the overall aim is to investigate the potential benefits of multimodality for pronunciation learning. In the current study, we explored L2 learners' ability to integrate auditory and textual input (i.e., spoken and written word forms) while reading dynamic texts (i.e., using subtitles) in L2-captioned video by temporally relating learners' eye fixations on words to their auditory onset. Previous research using eye-tracking measures to capture on-screen reading behavior suggests subtitled text benefits listening comprehension (Montero Perez *et al.,* 2013) and the incidental acquisition of L2 vocabulary (Bisson *et al*., 2014), but no research to date has examined the role of text-audio synchronization during bimodal input exposure.

Our working assumption is that more effective integration skills are likely to enhance gains in L2 speech perception and production over time via exposure to bimodal input. In addition, we assessed the contribution of individual differences in speech processing skills to L2 learners' processing and integration of auditory and textual input in dynamic texts. We addressed the following research questions:

1. To what extent can L2 learners integrate text and audio during bimodal input exposure through L2-captioned video?

2. Do individual differences in speech processing skills affect the effectiveness of text and audio processing during bimodal input exposure through L2-captioned video?

In order to address these questions it is crucial to assess the learners' L2 proficiency level. Given that L2 speech and text processing is less efficient in the L2 than it is in the L1 (Segalowitz, 2010), watching L2 captioned video is a particularly challenging activity for the L2 learner (especially at low proficiency levels) from a cognitive load perspective (Mayer, 2009). This is because multimodal processing requires resolving competition and integration of different sources of L2 input without having control over the speed of the information flow: the auditory input (the soundtrack), the action in the background (the scene), and the textual input (the subtitles). Under such circumstances, the L2 proficiency of learners as well as the use of their cognitive resources and speech processing skills are likely to contribute

substantially to their language processing efficiency and consequently to how much they can benefit from exposure to L2 captioned video for L2 pronunciation development.

METHODOLOGY

In order to assess L2 learners' processing of dynamic texts, we tracked their eye movements while watching short excerpts from the TV series *Sherlock* (2010) subtitled in English. The selected excerpts (1.5 minutes) featured quiet indoor conversations between characters without action in the background. In addition, we assessed L2 learners' individual differences in speech processing skills by obtaining accuracy measures from a speech processing test battery that included tasks on L2 speech segmentation, statistical learning of sound sequences, and modality integration.

Participants

Thirty-eight L1-Spanish/Catalan learners of English participated in the study for course credit. We assessed their proficiency level in English through an elicited imitation task (Ortega et al., 2002) consisting of 30 sentences with high-frequency vocabulary items, ranging 7-19 syllables in length, and of increasing grammatical complexity (see appendix). Participants heard each sentence only once through headphones and were asked to repeat the sentences as accurately as they could after a 2-second delay signaled by a beep sound. Following the scoring rubric of Ortega et al. (2002) available in the IRIS digital repository (Marsden, Mackey, & Plonsky, 2016), the first author scored each sentence assigning 0, 1, 2, 3 or 4 points, depending on how much of the sentence could be repeated and the kind of errors produced (if any), to a maximum score of 120 for the 30 sentences presented. The scores obtained ranged from 30 to 118 (M=96.87; SD=18.86), suggesting a relatively advanced level of proficiency at the high end of the scores.

Eye-tracking measures

Participants' eye movements were tracked and recorded on a Tobii T120 eye-tracker while they watched seven short clips. They were told they would do a language comprehension task that required watching the clips carefully in order to answer a true/false comprehension question appearing on the screen at the end of each one of the clips. The questions, which posed no difficulty to participants and were expected to be answered correctly, served to maintain students' attention and engagement while watching. For the current study, the eye-gaze data corresponding to two of the seven clips (Table 1) were used in the computation of the eye-tracking measures, as we estimated these data to be sufficient to provide individual measures of reading behavior.

Table 1

Characteristics of the clips

Clip	Duration (sec)	Subtitles (*n*)	Words (*n*)	Subtitle length (words)	Word length (characters)
1	89	33	220	6.3	5.3
2	95	35	271	7.9	4.9

Reading Index of Dynamic Text (RIDT)

The RIDT assesses the amount of visual processing of dynamic text. It provides a reliable 0-1 index of the visual processing of text in subtitles (Kruger & Stein, 2013). It is based on the number of fixations (points where the eye rests while reading) on each word in every caption while penalizing for skipping, re-fixations and regressions (Figure 1). The more text a viewer processes, the larger the index score, so that a viewer fixating all the words in all captions would obtain a score approaching 1. We used the RIDT to determine how much text in the captions learners had processed.

$$\text{RIDT} = \frac{\text{number of unique fixations for } p \text{ in } s}{\text{number of standard words in } s} \text{ x } \frac{\text{average forward saccade length for } p \text{ in } s}{\text{standard word length for } v}$$

(s=subtitle; p=participant; v=video)

Figure 1. Reading Index for Dynamic Text (RIDT) formula

Audio-Text Synchronization

The extent to which a fixation on a word in a caption is synchronized with its auditory presentation may determine how effectively the viewer can map auditory input to linguistic form. This mapping may be crucial in pronunciation learning in that it may promote changes in the phonological form of the learners' L2 lexical representations.

In order to estimate the extent to which auditory and visual word forms are synchronized while watching captioned video we determined whether participants fixated on the written form of a selected set of target words either *before* or *after* the occurrence of its auditory form. We also calculated the time distance in milliseconds between the onset of the fixation on the written word and the onset of the auditory form of the word in the sound track. Ten target words were selected on the basis of their length and position in the sentence. Words shorter than five characters and those appearing at the beginning or at the end of the sentence in the caption were avoided, as they are often skipped, according to research on caption reading (d'Ydewalle & De Bruycker, 2007)**.**

Speech processing measures

In what follows we describe the tasks used to obtain measures of participants' individual differences in speech processing skills. We obtained measures of speech segmentation (word-spotting and statistical learning) and audio-text integration (modality integration tasks). The elicited imitation task (described in the participants section) provided a measure of L2 proficiency.

Word-spotting

Participants' L2 segmentation skills were assessed through a word-spotting task (Cutler & Shanley, 2010; McQueen, 1996). Participants identified real English words embedded in nonwords by saying them out loud. We asked a male and a female native speaker of British English to produce 72 nonwords from the nonword set created by Farrell (2015), which were recorded in a sound-proof booth. The position of the real word in the nonwords were *final* (n=36, e.g. *ke*+*song*=ke*song*) or *initial* (n=36, e.g. *pound*+*fisp*=*pound*fisp). Nonwords conformed to one of three conditions: (1) *Stress*: The nonsense syllable was unstressed (n=24, e.g. *map*ef); (2) *Easy*: the segmentation affected an illegal phonotactic sequence in English (n=24, e.g. *ink*fab) or (3) *Difficult*: a legal one in English (n=24, e.g. *step*lut). Nonwords were presented auditorily only and participants had 3.5 seconds to say the embedded word, which was recorded to compute a percent correct accuracy score. Learners

with higher word-spotting scores were assumed to have better segmentation skills, which would provide them with an advantage in processing multimodal input.

Statistical learning

Individual differences in learners' ability to extract phonotactic regularities from sound sequences, an ability that supports the segmentation of words from fluent speech (Saffran, Aslin & Newport, 1996), were assessed through the statistical learning task in Palmer and Mattys (2016; stream A), based on Saffran et al. (1996). Participants were first exposed to a continuous stream of speech (6 minutes) consisting of 4 alien words (*laso*kachu, re*bufi*, *pema*dovi, ti*nugo*) and 4 part-words made up of word-final and word-initial syllables (*bufilaso*, *dovire*, *nugopema*, *kachuti*), each presented 125 times at a speech rate of 4.17 syllables per second. There were no acoustic segmentation cues within the speech stream, so that the relative between-syllable transitional probabilities within and across words was the main cue participants could use for inferring word boundaries. Then they performed a 24-trial recognition test where they were asked to identify which of two "alien" words, a possible one (one of the words or part-words) and a "nonword" which could not be made up of word-final and word-initial syllables (e.g., *finukado*) was a word in the alien language. Word pairs were presented auditorily and orthographically (500-ms ISI, 1000-ms ITI). We calculated an overall percent correct score based on the number of correctly identified words and part-words.

Modality integration tasks

Two modality integration tasks were designed to assess learners' skill at integrating text and sound in sentence-like contexts. Both tasks required learners to uncover text-sound mismatches as they simultaneously listened to and read short sentences appearing on the screen, half of which contained a text-sound mismatch. Their task was to decide whether what they heard matched what they read by pressing a designated *same* or *different* key on the keyboard. One task was in English (participants' L2) and the other one in a language participants could not understand (Basque). Thus, whereas in the English task participants had to process spoken and written word forms and their meaning, in the Basque task the processing of spoken and written word forms was based on phonetic and orthographic decoding only.

The English task contained two blocks of 24 sentence trials each produced by a female native speaker of Standard Southern British English (with 5 practice trials). In block A, the written sentences (adapted from Kennedy & Trofimovich, 2008) were presented in standard form. In block B, they were presented without spaces. In each block, half of the sentences were *same* trials (12) and half were *different* trials (12), and half of each one of these (6) contained a target nonword with an orthographic representation that either matched or mismatched its auditory form. We computed a percent correct identification score based on the number of correctly identified *same* and *different* trials.

The Basque task contained a single block of 20 sentence trials presented orthographically with spaces and produced by a female native speaker of Basque (with 4 practice trials), half of which (10) were *different* trials with text-sound mismatches. Learners who were better able to identify text-audio mismatches were assumed to show better text-audio integration skills, which would allow them to process multimodal input more effectively.

RESULTS AND DISCUSSION

Eye-tracking measures

The analysis of the eye-tracking data revealed a wide range of RIDT scores (Table 2), suggesting that the amount of text participants processed in the captions varied considerably across learners (from 0.16 to 0.78 within the 0-1 RIDT index). We were expecting the RIDT

score to correlate negatively with learners' proficiency, i.e. higher proficiency learners were expected to skip more words and captions than lower proficiency learners, as they would need to rely less on the supporting text in the captions to understand spoken language. In general, irrespective of L2 proficiency, we expected learners with more efficient speech processing skills to experience less difficulty in managing different sources of information and in integrating text and sound more efficiently.

Table 2

Eye-tracking measures (n= 38)

Measure	*M*	*SD*	*Min.*	*Max.*
RIDT	0.52	0.18	0.16	0.78
% words pre-fixated	70.4	17.2	33.3	100
% words post-fixated	29.6	17.2	0	66.7
Pre-fixation distance (milliseconds)	216	198	0	842
Post-fixation distance (milliseconds)	531	185	108	1183
Fixation distance (milliseconds)	441	120	186	760

The text-sound synchronization measures revealed that most fixations (70.4%) occurred before the auditory presentation of the word (Figure 2), with a large number of participants (71.3%) mainly fixating on the selected words before they were presented auditorily. This indicates, as expected, that participants had generally already read the selected target words by the time these occurred in the soundtrack. Such pre-fixations took place on average about 200 milliseconds from the onset of the auditory word form, a text-to-sound distance that was significantly shorter (Wilkoxon: T=680, z=4.48, p<.001) than the average time distance of about 500 milliseconds between the auditory word form and the following post-fixation (Figure 2). For the set of selected words in this analysis, the tendency was therefore for the auditory word form to follow (rather than precede) the lexical activation of the word via the visual text input.

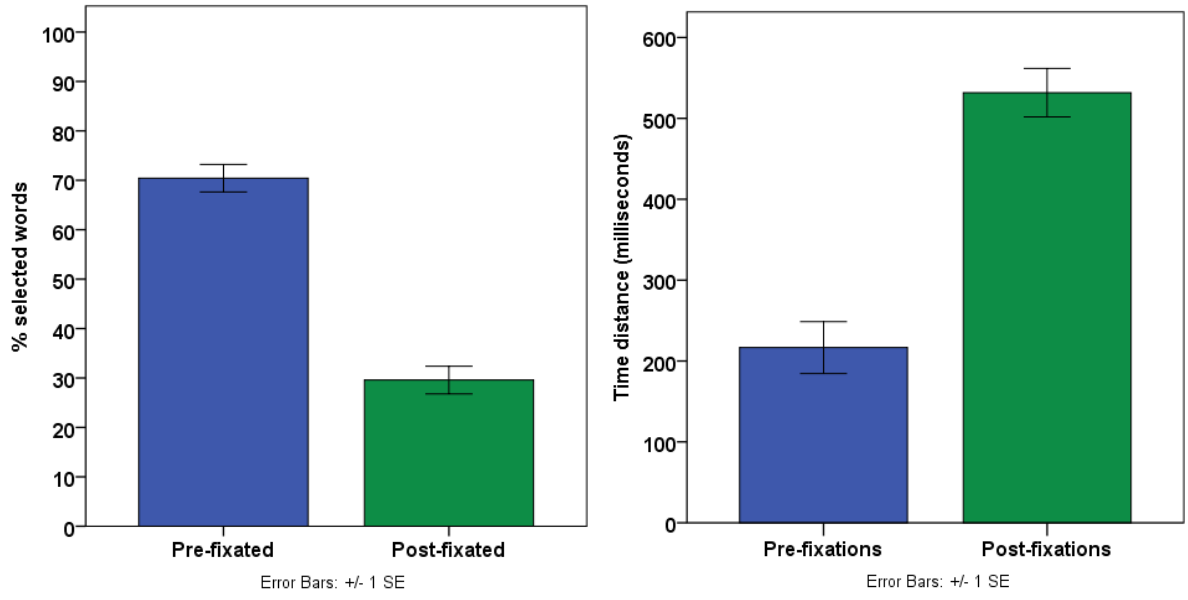


Figure 2. Percentage of selected words that were fixated before (pre-fixated) or after (post-fixated) their auditory onset (left) and distance in time between the eye-fixation on the word in the caption and its auditory onset (right).

Speech processing measures

The results obtained for the speech processing tasks (Table 3) also revealed large variability in L2 learners' performance both for tasks conducted in the L2 (reflecting inter-learner differences in L2 proficiency) and tasks conducted in a language unfamiliar to participants (L0).

Table 3

Descriptive data (% correct) for the speech processing tasks

Task	*M*	*SD*	*Min.*	*Max.*
Elicited imitation (L2)	56.7	9.7	41.7	81.9
Word-spotting (L2)	56.7	9.7	41.7	81.9
Statistical learning (L0)	56.6	15.9	25	96
English modality integration (L2)	81.6	9.4	55	95
Basque modality integration (L0)	75.1	9.4	55	95

The results of the elicited imitation task suggest an average upper intermediate level of proficiency, with scores ranging from intermediate to advanced proficiency levels.

Segmentation skills measured through the word spotting task, showed, as expected, differences in performance as a function of the *position* (initial, final) of the target word and *stimuli set* (stress, easy, difficult; see Figure 3). A two-way ANOVA with *position* and *stimuli set* as within-subjects factors revealed a non-significant main effect of *position* ($F(1, 36)=.230$, $p=.634$, $\eta^2=006$), a significant main effect of stimuli set ($F(2, 35)=20.01$, $p<.001$, $\eta^2=533$) and a significant *position* x *stimuli set* interaction ($F(2, 35)=36.33$, $p<.001$, $\eta^2=675$). As shown in Figure 2, the interaction arose because it was significantly harder to identify the target word in initial than in final position in the *stress* condition ($t(36)=-5.61$, $p<.001$), whereas the opposite happened in the *difficult* condition ($t(36)=3.98$, $p<.001$). The main effect of condition, however, was significant for both *initial* ($F(2, 35)=13.77$, $p<.001$, $\eta^2=440$) and *final* ($F(2, 35)=27.99$, $p<.001$, $\eta^2=615$) positions. In order to explore the relationship between segmentation skills and the eye-tracking measures (Table 3), we computed an average segmentation skill score across conditions by subject.

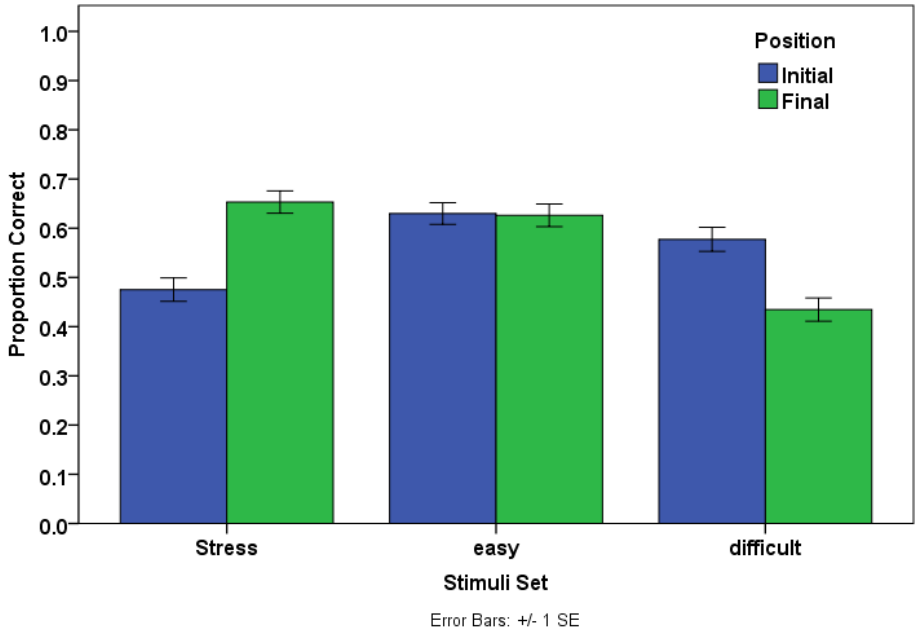


Figure 3. Proportion of correctly identified words by position and stimuli set.

The testing phase in the statistical learning task involved presenting participants with three sets of pairs of "alien" words for identification: *words* vs. *partwords*, *words* vs. *nonwords*, and *partwords* vs. *nonwords*. They were expected to identify *words* and *partwords* as words in the "alien" language at higher frequency rates than *nonwords*. As shown in Figure 4, mean scores by condition ranged between 50% and 60% (see Table 3). However, mean scores across testing conditions ranged from 25% to 96%, indicating large inter-learner variation, with 26 out of the 36 participants obtaining scores ≥ 50%. We computed a statistical learning score (Table 3) based on the average across the three testing conditions by subject.

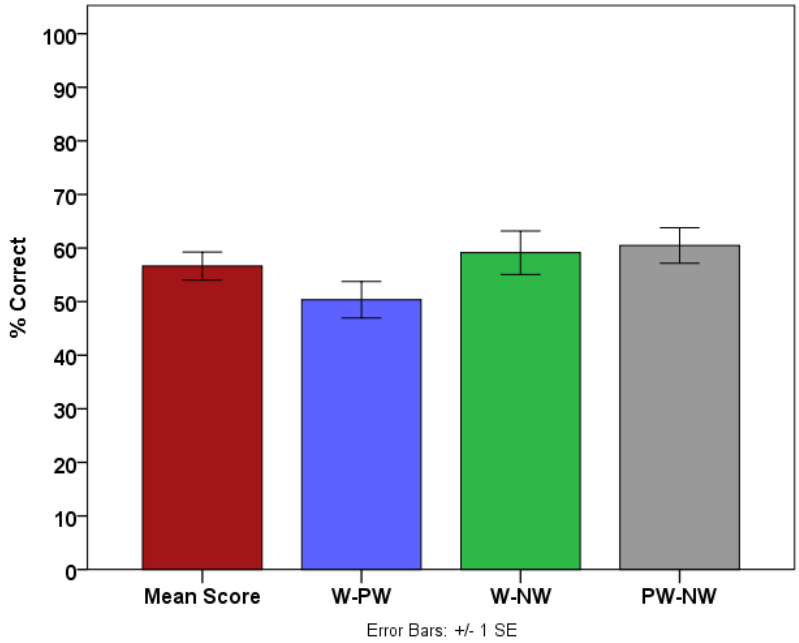


Figure 4. Proportion of correctly identified words (W) and partwords (PW).

Finally, as shown in Table 3, L2 learners' performance on the modality integration tasks also presented large inter-learner variation. Participants obtained significantly higher correct identification scores in the English than they did in the Basque version of the task ($t(37)=-3.42$, $p=.002$), suggesting that it was easier for them to identify mismatches when meaning processing was involved. Scores on both tasks appeared to be unrelated (*Pearson-r*=.141, $p=.383$), suggesting that they were tapping on different types of modality integration.

Correlation analyses

The main aim of the current study was to assess L2 learners' ability to integrate text and audio during bimodal input exposure through L2 caption video and to explore the role of individual differences in speech processing skills at doing so effectively.

We first explored whether our L2 proficiency measure was related to learners' performance in the L2 speech processing tasks. As expected, higher proficiency learners were better able to identify target words in the word-spotting task ($r=.446$, $p=006$), suggesting they could segment L2 speech better. They were also better able to detect text-sound mismatches in the English modality integration task ($r=.513$, $p<.001$), but not in the Basque modality integration task ($r=-.030$, $p=856$), suggesting that they were also better able to integrate text and sound in the L2. Proficiency scores, however, were found to be unrelated to any of the eye-tracking measures. Similarly, the L2 speech processing measures (word-spotting and English modality integration scores) were found to be unrelated to eye-tracking measures when L2 proficiency was controlled for. L0 speech processing measures (statistical learning and English modality integration scores) were unrelated to eye-tracking measures. However, as expected, higher RIDT indices were associated with the percentage of target words learners fixated on ($r=.552$, $p<.001$), suggesting that the RIDT index could reliably capture the amount of text learners processed in the captions. Finally, learners who were better able to identify text-sound mismatches in the Basque modality integration task had shorter time spans between auditory word onsets and fixations in general ($r=-.327$, $p=.045$), but larger time spans in post-fixations ($r=.502$, $p=.001$). This may be due to the fact that only a small proportion of fixations occurred after the auditory presentation of the word (only 29.6% were post-fixations), or it may indicate that text-sound synchronization skills are driven by the ability to read fast, suggesting that only "fast readers" can synchronize text and sound well.

CONCLUSION

This study was a first attempt at exploring the role of speech processing skills in bimodal input processing within our wider research aim (pronunciation development through L2 captioned videos). The results underscore the role of L2 proficiency in L2 speech processing and shed light on the mechanisms underlying text-sound synchronization during bimodal input processing in captioned video. Learners with higher L2 proficiency showed better L2 segmentation skills. However, whereas such skills were unrelated to eye-tracking measures, better L0 text-sound integration skills were related to longer time spans between auditory input and post-fixations, but not between auditory input and pre-fixations. This may suggest an important role for individual differences in reading speed (which we did not test) in text-sound integration when watching captioned video. Testing L1 and L2 silent reading speed and eye-gaze behavior is needed in order to better understand the mechanisms underlying text-sound integration during exposure to bimodal input and further explore their role in L2 pronunciation development. Our next step in this research project is to relate the measures explored in the current study to gains in L2 pronunciation obtained after extensive exposure to L2-captioned videos.

ACKNOWLEDGMENTS

We would like to thank Ingrid Mora-Plaza and Diana Moreira de Oliveira for their help in data collection.

ABOUT THE AUTHORS

Joan C. Mora is associate professor in the Department of Modern Languages and Literatures and English Studies in University of Barcelona (UB). His research has examined the role of contextual and individual factors in the development of L2 speech and oral fluency, and the acquisition of L2 phonology.

Contact information: mora@ub.edu, https://joancmora.weebly.com/

Natalia Wisniewska is a PhD candidate and research assistant collaborating with the Language Acquisition Research Group (GRAL) at the University of Barcelona. Her PhD dissertation focuses on the potential benefits of multimodal input for pronunciation development and the role of individual cognitive differences in the acquisition of L2 speech.

Contact information: wisniewska@ub.edu

REFERENCES

Bisson, M.J., Van Heuven, W.J., Conklin, K. and Tunney, R.J. (2014) Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics,* 35 (2), 399–418.

Broersma, M. (2012). Increased lexical activation and reduced competition in second-language listening. *Language and Cognitive Processes*, 27, 1205-1224.

Charles, T. J., & Trenkic, D. (2015). Speech segmentation in a second language: The role of bimodal input. In Y. Gambier, A. Caimi, & C. Mariotti (Eds.), *Subtitles and Language Learning: Principles, Strategies and Practical Experiences,* (pp. 173-198). Bern: Peter Lang.

Danan, M. (2004). Captioning and subtitling, undervalued language learning strategies. *Meta*, *49*(1), 67–77.

d'Ydewalle, G., & De Bruycker, W. (2007). Eye movements of children and adults while reading television subtitles. *European Psychologist, 12*(3), 196-205. DOI: 10.1027/10169040.12.3.196.

Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility, and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, *64*(3), 459-489.

Kruger, J. L., & Steyn, F. (2014). Subtitles and eye tracking: Reading and performance. *Reading Research Quarterly*, *49*(1), 105-120.

Marsden, E., Mackey A., & Plonsky, L. (2016). The IRIS Repository: Advancing research practice and methodology. In A. Mackey & E. Marsden (Eds.), *Advancing methodology and practice: The IRIS Repository of Instruments for Research into Second Languages* (pp. 1-21). New York: Routledge.

Mayer, R. E., (2009). *Multimedia Learning*. Second Edition. New York: Cambridge University Press.

McQueen, J. (1996). Word spotting. *Language and Cognitive Processes*, *11*(6), 695-699.

McQueen, J. M., & Cutler, A. (1998). Spotting (different kinds of) words in (different kinds of) context. In R. Mannell, & J. Robert-Ribes (Eds.), *Proceedings of the Fifth International Conference on Spoken Language Processing: Vol. 6* (pp. 2791-2794). Sydney: ICSLP.

Mitterer, H., & McQueen, J. M. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PloS one*,*4*(11), 1-5.

Montero Perez, M., Van Den Noortgate, W., & Desmet, P. (2013). Captioned video for L2 listening and vocabulary learning: A meta-analysis. *System*, 41(3), 720-739.

Mora, J. C., Levkina, M., (2017). Task-based pronunciation teaching and research: key issues and future directions. *Studies in Second language Acquisition*, 39, 381-399.

Munro, M. J., Derwing, T. M., & Thomson, R. I. (2015). Setting segmental priorities for English learners: Evidence from a longitudinal study. *International Review of Applied Linguistics in Language Teaching*, *53*(1), 39-60.

Ortega, L., Iwashita, N., Norris, J. M., & Rabie, S. (2002, October). An investigation of elicited imitation tasks in crosslinguistic SLA research. Presentation given at *Second Language Research Forum,* Toronto.

Paivio, A. (1986). *Mental representations, A dual coding approach*. Oxford: Oxford University Press.

Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *The Quarterly Journal of Experimental Psychology, 69*(12), 2390-2401.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 1926-1928.

Segalowitz, N. (2010). *Cognitive Bases of Second Language Fluency*. New York: Routledge.

Sherlock (2010) Season 1, Episode 1: *A Study in Pink*, episode 2: *The Blind Banker*, TV programme. Directed by P. McGuigan. UK: BBC.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*,36(3), 326-344

Vanderplank, R. (1988). The value of teletext subtitles in language learning. *ELT Journal, 42*(4)*,* 272–281.

APPENDIX

The sentences in the elicited imitation task (from Ortega et al., 2002) available from IRIS (Marsden, Mackey, & Plonsky, 2016)

1. I have to get a haircut.
2. The red book is on the table.
3. The streets in this city are wide.
4. He takes a shower every morning.
5. What did you say you were doing today?
6. I doubt that he knows how to drive that well.
7. After dinner I had a ling, peaceful nap.
8. It is possible that it will rain tomorrow.
9. I enjoy movies which have a happy ending.
10. The houses are very nice but too expensive.
11. The little boy whose kitten died yesterday is sad.
12. That restaurant is supposed to have very good food.
13. I want a nice, big house in which my animals can live.
14. You really enjoy listening to country music, don't you?
15. She just finished painting the inside of her apartment.
16. Cross the street at the light and then just continue straight ahead.
17. The person I'm dating has a wonderful sense of humor.
18. She only orders meat dishes and never eats vegetables.
19. I wish the price of town houses would become affordable.
20. I hope it will get warmer sooner this year than it did last year.
21. A good friend of mine always takes care of my neighbor’s three children.
22. The black cat that you fed yesterday was the one chased by the dog.
23. Before he can go outside, he has to finish cleaning his room.
24. The most fun I've ever had was when we went to the opera.
25. The terrible thief whom the police caught was very tall and thin.
26. Would you be so kind as to hand me the book which is on the table?
27. The number of people who smoke cigars is increasing every year.
28. I don't know if the 11:30 train has left the station yet.
29. The exam wasn't nearly as difficult as you told me it would be.
30. There are a lot of people who don’t eat anything at all in the morning.

Levis, J. M. & Silpachai, A. O. (2018). Prominence and information structure in pronunciation teaching materials. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 216-229). Ames, IA: Iowa State University.

PROMINENCE AND INFORMATION STRUCTURE IN PRONUNCIATION TEACHING MATERIALS

John M. Levis, Iowa State University
Alif O. Silpachai, Iowa State University

Prominence is marking of particular syllables as salient in English speech. This marking is accomplished by the pitch, duration and intensity of the voice, and is multi-functional in English. Prominence is the target of increasing research both in regard to its form and its functions Prominence is also one of the most commonly taught suprasegmental features included in published pronunciation materials, and it is uniformly seen by pronunciation researchers as critical to intelligibility. The linguistic and pedagogical research on prominence, however, has diverged, and very little theoretical research is reflected in pronunciation teaching materials. This paper examines what current research shows about the form and functions of prominence in English, describes how prominence is represented in teaching materials, and suggests areas of current research that can profitably be applied to teaching materials.

INTRODUCTION

One of the most commonly taught suprasegmental features of English pronunciation is prominence (also known as sentence stress, nuclear stress, tonic, etc.). Prominence is not only commonly taught in pronunciation materials, it is also the subject of a wide range of current research in regard to both its form and its functions. In regard to form, prominence is the use of pitch, duration, and intensity to mark particular words/syllables in an utterance as salient. Functionally, prominence has multiple uses, the most important of which are to mark a default placement on the final content word of a phrase, to mark contrasting information, and to signal new information and given information. The purpose of this paper was to examine how pronunciation teaching materials reflect the findings of linguistic research on prominence and to suggest possible changes to teaching materials to connect them more closely to current findings.

PROMINENCE: FORM, FUNCTION AND PERCEPTION

Prominence – Its form

Prominence refers to the greater strength of a word or a syllable compared to other words or syllables surrounding it within a phonological phrase. In English, for example, some prominent syllables are perceived as more important than others, and they often bear stress accents (Beckman, 1986). Prominence in English can be phonetically marked in many ways. The most common acoustic cues to prominence are fundamental frequency (f0), duration, intensity, segmental clarity, and any combination of these features.

Particularly, prominent words often have salient f0 movements expressing pitch accents (Gussenhoven, Repp, Rietveld, Rump, & Terken, 1997; Ladd 1996; Pierrehumbert 1980; Rietveld & Gussenhoven, 1985; Terken, 1991), increased duration and/or intensity, increased spectral emphasis in the mid and high frequency regions relative to non-prominent words (Beckman 1986; Beckman & Edwards 1994; Cambier-Langeveld & Turk 1999; Cole, Kim, Choi, & Hasegawa-Johnson, 2007; Kochanski, Grabe, Coleman, & Rosner, 2005; Sluijter & van Heuven 1996; Tamburini 2005; Turk & White 1999). Prominent words are also often hyper-articulated, relative to non-prominent words. That is, these words are pronounced more clearly than usual, and as a result they have larger vowel spaces (Baker & Bradlow, 2009).

Prominence – Its functions

Prominence at the phrasal level is often identified with the information structure of the phrase. Specifically, prominent words often introduce information that is new or important to the goal of the discourse, or they may bear contrastive focus (Bolinger 1986). In contrast, words that lack prominence are typically considered given in the prior discourse, or anaphorically recoverable (Schwarzchild, 1999). The relationship between prominence and information structure is typically strong in rightmost prominent words (words that bear nuclear accents) in the phrase (Calhoun, 2006), whereas prominence in pre-nuclear positions seems to depend on other factors, such as those that affect rhythm (Cole, Mo, & Hasegawa-Johnson, 2010).

Prominence – Its perception

There are conflicting answers about which phonetics cues reliably mark prominence. Early perceptual studies of single words by Fry (1955, 1958) suggest that prominent syllables are marked, in decreasing order of importance, by duration, f0, and intensity. Using analyses from laboratory speech, Lieberman (1960) described a system for deducing lexical stress from acoustics. His work suggests that these three cues are similarly important; that is, each cue is a good predictor of prominence. Other studies using laboratory speech have yielded varying results. For example, perceptual studies by Gussenhoven et al. (1997), Rietveld and Gussenhoven, (1985), and Terken, (1991) found that f0 bumps in synthesized words were perceived as prominent. Beckman (1986) found that prominence substantially correlated with a combination of intensity and duration. Lastly, the synthesis experiments by Turk and Sawusch (1996) suggested that duration and intensity are perceived together as a single percept, although the results of their rating scale experiment indicated that intensity does not significantly contribute to perceived prominence.

Experiments using natural speech (e.g., spontaneous speech) have found that f0 plays a relatively minor role in prominence perception. For example, in Silipo and Greenberg (1999, 2000), two trained linguists agreed, when asked to manually mark prosodic stress in spontaneous American English discourse, that intensity and duration played the major role in marking prominence and that pitch of vocalic nuclei played only a minor role. Similar results were found in another corpus study by Kochanski et al. (2005) which examined how prominence is acoustically marked in speech in a database covering several dialects of British and Irish English and three speech styles. It was found that speakers generally did not use f0 to distinguish prominent syllables from other syllables

within an utterance. Instead, prominence was primarily marked with intensity and duration.

Acoustic features related to prominence perception interact with other factors related to pragmatics and discourse, making it difficult to precisely specify how prominence is phonetically marked. For example, words from sparse lexical neighborhoods are often phonetically reduced compared to those from dense lexical neighborhoods (Munson, 2007; Wright, 2004). Moreover, words preceded by a highly probable context are more phonetically reduced than the same words in a less probable context (Lieberman, 1963). The classic example is the word *nine* in *a stitch in time saves nine* and *the number that you will hear is nine*, where the target word is preceded by a more probable context in the first sentence. Third, high frequency words in a language tend to have less salient f0 marking, reduced duration and intensity, and decreased vowel formant dispersions relative to low frequency words (Aylett & Turk 2004; Bell, Jurafsky, Fosler-Lussier, Girand, & Gregory, 2003; Fossler-Lussier & Morgan 1999; Gregory 2002; Ito, Speer & Beckman 2004; Munson 2007; Watson, Arnold & Tanenhaus 2008; Wright 2004).

Furthermore, a word's phonetic realization tends to be reduced on its second or subsequent mention (Baker & Bradlow, 2009; Fowler & Housum, 1987). While earlier studies suggest that second mention reduction might be induced by a word's discourse-given status (Bard, Lowe, & Altmann, 1989), recent research (Baker & Bradlow, 2009) suggests that the second mention reduction may also occur when the apparent second mention does not have the same referent as the first mention, indicating that second mention reduction is not purely semantically motivated. The effects of lexical frequency and previous mention may not only occur in the acoustic signal but also in the listener's mind. Cole et al. (2010) found that listeners tended to rate low frequency words as prominent even when these words lacked the necessary acoustic cues for prominence (in their study, increased duration and intensity).

INFORMATION STRUCTURE IN TEACHING MATERIALS

In English pronunciation teaching materials, prominence is typically presented as a required element of prosody. The form involves marking a syllable in a phrase as more prominent than other syllables. Prominent syllables are typically said to occur once (and sometimes more than once) within each spoken phrase. In regard to pitch, a prominent syllable is usually represented with a pitch excursion up or down from the pitch line, as in (1) and (2). In (1), the stressed syllable of the final word has a jump in pitch (the prominent syllable). This is followed a fall in pitch to the end of the sentence. In (2), the utterance has the same prominent syllable, but it starts at a relatively low pitch before rising to a high pitch on the last syllable.

(1) I'm going to Argen **TI** n a.

(2) You're going to Argen **TI** n a?

In (1) and (2), prominence is in the default position for English, that is, on the stressed syllable of the last content word in the phrase. Up to 90% of English phrases in spontaneous spoken language have prominence in this position (Crystal, 1969). Prominence placement may deviate from the default position in a number of sentence structures because some information is not expressed in the phrase. In (3), prominence (marked in CAPS) is not on the final content word because the sentence ends with a time-adverbial (Allerton & Cruttenden, 1979; Dickerson, 1989).

(3) He's GOing soon.

Perhaps the most commonly taught non-final placement of prominence is when prominence signals the information structure of discourse. This function of prominence includes two aspects of the system. First, a word or syllable is marked as prominent because it is new information. Second, and equally important, final content words may be non-prominent when they are no longer new, that is, when they are given.

The identification of new and given information is typically presented as being straightforward, with lexical items that were previously new being repeated (and so becoming given). Other lexical items that were not previously in the discourse then presented as new, and are thus identifiable through their prominence. An example of this is found in (4), from Grant (2012, p. 114).

(4) A. Let's continue our discussion of polLUtion. / B. YESterday / C. we deFINED pollution. / D. ToDAY / E. we'll talk about the IMpact of pollution / F. its far-reaching efFECTS. / G. Many people think pollution is just a problem for SCIentists / H. but it's NOT just a problem for scientists. / I. It affects EVeryone. / J. Because it affects human LIVES, / K. it's a HEALTH problem. / L. Because it affects PROperty, / M. it's an ecoNOmic problem. / N. And because it affects out appreciation of NAture, / O. it's an aesTHEtic problem."

In (4), we see a constructed paragraph to show how prominence (in CAPITAL letters) highlights new information and how lack of prominence on final content words can signal that the lexical item can mark information as given. For example, the word *pollution* is marked with prominence in phrase A. In A, *pollution* is phrase final and there is no reason to mark anything else as prominent because it is the first phrase. In C and E, *pollution* is again phrase final but is not prominent. *Pollution* has become given information, and prominence marks the new information, the next to last content word (*deFINED* and *IMpact*). Another example of a lexical item starting as prominent and then becoming non-prominent is the word *PROBlem* in phrase G. The word is repeated in H, K, M and O, three times as the last content word. But in each case, *problem* does not receive prominence because it is given information.

Prominence and information structure in pronunciation teaching materials

When considering texts like that found in (4), the Given/New distinction seems straightforward at first glance. The cognitive challenges of identifying Given/New in constructed texts is passed over, and the even greater cognitive challenges of making use of prominence to express information structure in spontaneous speech is almost never addressed. (Levis, 2001; Levis & Grant, 2003). The ways that L2 learners are taught about information structure raise several concerns.

- *Prominence is multi-functional in English and does not simply mark New and Given information*. Because prominence may also be used to call attention to contrasts (Levis & Muller Levis, 2018), to correct misinformation and to emphatically agree (Grant, 2012), L2 learners may struggle to distinguish other functions of prominence from prominence's role in marking information structure.
- *Information structure is not always as clear as constructed passages suggest*. New information placement overlaps with final content word placement because new information is often on the final content word due to grammatical elision, e.g., *I lost my umBRELla. What KIND?* (*What kind* has prominence on the final content word, but is short for *What kind of umbrella did you lose?* The missing words after *KIND* are understood from the original question.)
- *Lexical repetition does not always involve the same words, and lexical items that refer to the same thing are not always marked as given*. Information that is not lexically identical may be considered given because of its understood relationship to the original word (e.g., *Did you buy the CAtamaran? No, I had to get a SMALler boat*.) On the other hand, related words may be presented as different from the initial mention (e.g., *Have you even flow in an AIRplane? Sure. Last month, I went to Europe on a large JET.)*
- *Teaching students to recognize new and given information is difficult*. The cognitive aspects of such decisions, especially in longer or spontaneous texts, seems to assume native speaker competence in interpretation (Levis, 1999).
- *Teaching the pronunciation of information structure is easiest when using a prewritten text and when using clearly defined rules*. Although this type of pronunciation practice can be effective in the short run, it does not necessarily last (Hahn, 2002) or transfer to spontaneous speech.

As an illustration of how simplifying information structure for pronunciation teaching can actually make the topic quite complex, Table 1 shows how informational stress (Given-New) is taught in one pronunciation book (Reed & Michaud, 2005, p. 127). The explanation mixes several functions together and talks about words being prominent because they are important, an unexplained evaluation, rather than because they are final. The explanations also conflate the typical prominence on content words and the less common prominence on function words, the use of prominence for contrasts, and finally, conflates new information and contrast without explaining what is being contrasted.

When do you use informational stress?	
Dialogue	**Explanation**
A: Where's the book?	The content word "book" is the most important word in this question. It receives standard sentence-level stress.
B: The book's on the counter	The word "book" is now old information, so the content word "counter"-the new piece of information-is stressed (informational stress).
A: Next to the paper?	The content word "paper" is the most important word in this question. It's new information. Notice that the stressed words in all the examples so far have been content words. Usually the most important word in a sentence is a content word. However, this isn't always the case.
B: No, *under* the paper.	Here, the most important word is the preposition "under" (a function word) because it's a new piece of information and because it contrasts with "next to."
A: I've already looked under the paper.	The content word "looked" is the new piece of information. "Looked" receives informational stress.
B: Well, look *again*	Now the word "look" is old information. The word "again" is stressed because it is the new piece of information and it's contrastive.

The relationship of the default placement of prominence on the last content word, and the use of prominence to mark new information occurring on the last content word is sometimes addressed by pronunciation textbooks, but there is usually be no clear explanation about why the same prominence placement has two different explanations, as in the examples in (5) and (6) from Lane (2005, p. 166). The example also does not show new information that is not at the end of a sentence. This requires language learners and teachers to provide such information on their own. If they do not understand the system, however, this may prove impossible.

(5) Beginning a Conversation: When you begin a conversation, you often highlight the last content word.

What did you do on the WEEKend?

(6) Highlighting New Information: New information is often presented in the last content word of a sentence.

(What did you do on the WEEKend?) I went DANcing.

Some textbooks try to be more systematic in explaining new information, but in doing so they may increase cognitive complexity, as in the examples in (7)-(9) from Dauer (1993, p. 231). Dauer explains what is meant by new and old (given) information

Pronunciation involves both cognitive and procedural knowledge

SENTENCE STRESS ON NEW INFORMATION

Sentence stress is also moved to separate new information from old information. Old information is what the speaker assumes the listener already knows, either because it was just mentioned in a previous sentence or because it is part of the physical situation. Sentence stress will fall on the new information. If the old information is repeated, it will not receive sentence stress. In the following examples, the same meaning can also be expressed by using auxiliaries, omitting the old information, reordering the sentence, or using pronouns.

(7) A: Who borrowed my eraser?

B: I borrowed it. (== I did.)

I is new information, not known by A; *borrowed it* is old information.

(8) A: I bought a new car.

B: What **kind** of car did you buy? (== What kind?)

(9) Teacher: This is a difficult test. (== This test is difficult.)

The teacher has the test in her hands, so it's known or old information.

Simplifying to make information accessible often involves assumptions about whey certain types of lexical items are prominent while others are not, as in (10) from Gilbert (2012, p. 60) in which new information is described as marking a new thought, as though each lexical item represented a thought. Additionally, there is now clear statement of why *KIND* is new in B: but *of* is not.

(10) After a conversation begins, any word can become a new thought (the new focus of information).

A: I lost my HAT. ("Hat" is the last content word. It is the focus of the sentence.)

B: What KIND of hat? ("Kind" is now the focus. It is the new thought, and "hat" is an old thought.)

A: It was a RAIN hat. ("Rain" is now the focus. It is the new thought.)

In (11), the difficulty of representing new information can be seen in the use of words within the same lexical set, in this case, money and dollars, which is clearly a synonym for money within an American English context, yet is described as representing new information (Miller, 2000, p. 71).

(11) Use focus to highlight new information. Stress the word that gives the new information.

A: I need to borrow some **MO**ney. ("money" is new information)

B: How **MUCH** money? ("money" is now old information)

A: Well, not **TOO** much money. ("much" and "money" are both old information)

B: I have about ten **DOL**lars. ("dollars" is new information)

A: I was hoping to borrow **TWEN**ty dollars. ("dollars" is now old information)

SUGGESTED CHANGES FOR MATERIALS

Prominence is considered by pronunciation researchers as a critical feature for intelligibility (Hahn, 2004; Jenkins, 2000), especially in relation to prominence's function in marking information structure. However, current pronunciation materials, in their desire to make new and given information accessible to L2 learners, often simplify in ways that do not reflect what research tells us about prominence. In this section, we suggest directions for changes in pronunciation materials that can make current insights into prominence and its role in signaling information structure. Here we present four suggestions for connecting pronunciation teaching practices more closely to research.

1. Use real spoken data (and longer texts) to help learners perceive prominence in speech and to help learners work out patterns

This recommendation is to use not only constructed texts in teaching new and given information, but asks us to also make use of authentic spoken texts. L2 learners, especially at higher proficiency levels, can analyze such texts to cognitively engage with how speakers construct discourse and highlight particular words and syllables to communicate their message.

2. Describe how to identify "information", what makes something "new" or "given", and the relationship of new information to the default pattern

Information or thoughts are implicitly associated with particular lexical items in discourse, but materials often assume L2 learners will be able to apply example texts to

new texts and to spontaneous speech. This is not accurate. While the pronunciation of prominence may be quite teachable (Levis and Muller Levis, 2018; Pennington & Ellis, 2000), the cognitive aspects of prominence placement are far more difficult and cannot only be addressed in relation to perception and production.

3. Explicitly practice given/old information

Marking new information is often taught as the only important function of prominence, but equally important is the marking of given information. While new information is marked as phonologically salient, given information must be backgrounded both to avoid calling attention to it and to contrast with the salience of the prominent syllables. Our experience has been that L2 learners can mark words associated with new information as prominent but that following words associated with given information are also marked as prominent rather than being deaccented. Almost no pronunciation teaching materials explicitly teach deaccenting of given information despite its importance in the prosodic shape of an utterance.

4. Include exercises to fill out the communicative framework for teaching pronunciation (Celce-Murcia et al., 2010) to encourage moving beyond controlled production to cognitive understanding of information structure.

Pronunciation teaching involves varied activities and exercises to address the complex and interrelated skills involved in L2 learning. L2 learners, especially adult learners, need cognitively oriented explanations of the pronunciation feature and how it functions, practice hearing and interpreting the feature, and training and rehearsal in producing the feature with and without attention to communicative meaning. In Table 1, we use the five-part communicative teaching framework of Celce-Murcia, Brinton and Goodwin (2010) to suggest possible changes to the way we teach prominence and information structure.

Table 1

What is Currently Available in Textbooks and What is Needed

Communicative Pronunciation Teaching Stages	Evaluation of current materials
Explanation/Analysis	Current: Often inadequate, with poorly described rules. Needed: Descriptions of patterns that reflect to complexity of information structure, and how information structure is related to other functions of prominence.
Perception	Current: Simple listening only with no perception training (e.g., identify prominent words in a spoken text).

	Needed: Better perception activities using multiple voices. Use of authentic speech that demonstrates information structure. Discussion exercises allowing L2 learners to discuss and interpret meaning.
Controlled production (strong focus on intonation form, e.g., read aloud)	Current: This stage dominates practice activities related to prominence. Needed: More practice on marking prominence and deaccenting given information. Practice using delayed repetition exercises rather than simple reading aloud. The goal here is to build automaticity of production.
Guided production (some focus on meaning required along with some focus on intonation form, e.g., simple information gap activities)	Current: Rare in published materials Needed: More exercises that allow learners to practice the form of prominence while also paying attention to meaning in their practice. This would involve activities such as simple information gap and role play exercises.
Communicative production (focus on meaning dominates, e.g., discussion or debate)	Current: Rare in published materials Needed: Activities that allow free expression but that also require learners to express new and given information, such as presentations or debates. These can be recorded and used for analysis before repeating the activity.

CONCLUSION

The use of prominence to signal the information structure of discourse is a critical aspect of communicative ability in English. It is also a cognitively challenging aspect of speech for L2 learners who may not understand either the pronunciation or the functions of prominence. We suggest that more effective teaching of this feature must take into account non-pedagogical research on prominence and on information structure.

ABOUT THE AUTHORS

John Levis is Angela B. Pavitt Professor of English at Iowa State University. His articles on pronunciation and intonation have been published in a variety of professional journals, including *TESOL Quarterly, Applied Linguistics, System, Annual Review of Applied Linguistics, TESOL Journal, ELT Journal* and *World Englishes*. He is author of *Intelligibility, Oral Communication and the Teaching of Pronunciation* (Cambridge University Press). He was co-editor for the Phonetics and Phonology section of the *Encyclopedia of Applied Linguistics* (Blackwell), *Social Dynamics in Second Language Accent* (De Gruyter Mouton), *Handbook of English Pronunciation* (Wiley), and *Critical*

Concepts in Linguistics: Pronunciation (Routledge). He initiated the annual Pronunciation in Second Language Learning and Teaching Conference and is founding editor of the *Journal of Second Language Pronunciation* (John Benjamins). His newest project is *Pronunciationforteachers.com*, a website providing reliable information about teaching pronunciation. Email: jlevis@iastate.edu

Alif Silpachai is graduate student in Applied Linguistics and Technology at the Department of English at Iowa State University. His research interests include production and perception of suprasegmentals, particularly pitch accents and lexical tones. Email: alif@iastate.edu

REFERENCES

Allerton, D. J., & Cruttenden, A. (1979). Three reasons for accenting a definite subject. *Journal of Linguistics*, *15*(1), 49-53.

Arnold, J. E. (2008). THE BACON not the bacon: How children and adults understand accented and unaccented noun phrases. *Cognition*, *108*(1), 69-99.

Aylett, M., & Turk, A. E. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47(1), 31–56.

Baker, R. E., & Bradlow, A. R. (2009). Variability in word duration as a function of probability, speech style, and prosody. *Language and speech*, *52*(4), 391-413.

Baltazani, M. (2006). Intonation and pragmatic interpretation of negation in Greek. *Journal of Pragmatics*, *38*(10), 1658-1676.

Bard, E. G., Lowe, A. J., & Altmann, G. T. (1989). The effect of repetition on words in recorded dictations. In *First European Conference on Speech Communication and Technology* (pp. 2573-2576). Retrieved from https://www.isca-speech.org/archive/archive_papers/eurospeech_1989/e89_2573.pdf

Baumann, S., & Grice, M. (2006). The intonation of accessibility. *Journal of Pragmatics*, *38*(10), 1636-1657.

Beckman, M. E. (1986) *Stress and non-stress accent*. Dordrecht, Foris.

Beckman, M. E. & Edwards, J. (1994). Articulatory evidence for differentiating stress categories," in P. Keating (Ed.),*Phonological structure and phonetic form* (pp. 7-33). Cambridge: Cambridge University Press.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., & Gregory, M. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *The Journal of the Acoustical Society of America*, 113, 1001–1024.

Birch, S., & Clifton, C. (1995). Focus, accent, and argument structure: Effects on language comprehension. *Language and speech*, *38*(4), 365-391.

Bishop, J. (2012). Information structural expectations in the perception of prosodic prominence. In G. Elordieta & P. Prieto (Eds.), *Prosody and meaning (interface explorations). Berlin: Mouton de Gruyter*.

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

Calhoun, S. (2006). *Information structure and the prosodic structure of English: A probabilistic relationship* (Doctoral dissertation, University of Edinburgh, UK).

Cambier-Langeveld, T., & Turk, A. (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, 27, 255–280.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs (2): A course book and reference guide*. Cambridge University Press.

Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from radio news speech. *Journal of Phonetics*, 35, 180–209.

Cole, J., Mo, Y., & Hasegawa-Johnson, M. (2010). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology*, *1*(2), 425-452.

Crystal, D. (1976). *Prosodic systems and intonation in English*. Cambridge University Press.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*(2), 292-314.

Dauer, R. (1993). *Accurate English*. Englewood Cliffs, NJ: Regents/Prentice Hall

Dickerson, W. (1989). Stress in the stream of speech: The rhythm of spoken English. *Urbana, IL: University of Illinois Press*.

Fosler-Lussier, E., & Morgan, N. (1999). Effects of speaking rate and word frequency on pronunciations in convertional speech. *Speech Communication*, 29(2-4), 137-158.

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *The Journal of Memory and Language*, 26(5), 489-504.

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

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

Gilbert, J. (2012). *Clear speech*, 4th Ed. New York: Cambridge University Press.

Grant, L. (2012). *Well Said*, 3rd Ed. Boston: Heinle Cengage.

Gregory, M. (2002). Linguistic informativeness and speech production: An investigation of contextual and discourse pragmatic effects on phonological variation. (Doctoral dissertation, University of Colorado at Boulder).

Gussenhoven, C., Repp, B. H., Rietveld, A., Rump, H. H., & Terken, J. (1997). The perceptual prominence of fundamental frequency peaks. *The Journal of the Acoustical Society of America*, *102*(5), 3009-3022.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL quarterly*, *38*(2), 201-223.

Hahn, M. K. (2002). *The persistence of learned primary phrase stress patterns among learners of English* (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Ito, K., Speer, S. R., & Beckman, M. E. (2004). Informational status and pitch accent distribution in spontaneous dialogues in English. In *Speech Prosody 2004, International Conference*. Retrieved from https://www.isca-speech.org/archive_open/sp2004/sp04_279.pdf

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, *118*(2), 1038-1054.

Ladd, R. D. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Lane, L. (2005). *Focus on pronunciation 3*. White Plains, NY: Longman.

Levis, J. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly, 33*, 37-63.

Levis, J. (2001). Teaching focus for conversational use. *ELT Journal, 55*, 47-54.

Levis, J. (2018). *Intelligibility, oral communication and the teaching of pronunciation*. Cambridge University Press.

Levis, J., & Grant, L. (2003). Integrating pronunciation into ESL/EFL classrooms. *TESOL Journal 12*(2), 13-19

Levis, J. & Muller Levis, G. (2018). Teaching high value pronunciation features: Contrastive stress for intermediate learners. *CATESOL Journal 30*(1), 139-160.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*. 32, 451–454.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6(3), 172-187.

Miller, S. (2000). *Targeting pronunciation*. Boston: Houghton Mifflin.

Munson, B. (2007). Lexical access, lexical representation, and vowel production. In J. Cole & J. I. Hualde (Eds.). *Laboratory phonology* (Vol. 9, pp. 201–228). Berlin: Mouton de Gruyter.

Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal*, *84*(3), 372-389.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (Doctoral dissertation, Massachusetts Institute of Technology).

Prevost, S. A. (1996). *A semantics of contrast and information structure for specifying intonation in spoken language generation*. IRCS Technical Reports Series. 6.

Reed, M. & Michaud, C. (2005). *Sound concepts*. New York: McGraw Hill.

Rietveld, A. C. M. and Gussenhoven, C. (1985). On the relation between pitch excursions and prominence. *Journal of Phonetics*, *13*, 299–308.

Schwarzchild, R. (1999). GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics*, 7, 141 177.

Silipo, R. and Greenberg, S. (1999). Automatic transcription of prosodic stress for spontaneous English discourse. In Proceedings of the XIVth International Congress of Phonetic Sciences ICPhS99 , pp. 2351–2354.

Silipo, R. and Greenberg, S. (2000). Prosodic stress revisited: Reassessing the role of fundamental frequency. In Proceedings of the NIST Speech Transcription Workshop. College Park, MD. Retrieved from http://http.icsi.berkeley.edu/ftp/global/global/pub/speech/papers/niststw00-stress.pdf

Sluijter, A. M. C., & van Heuven, V. (1996). Spectral tilt as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, 100, 2471–2485. doi:10.1121/ 1.417955

Tamburini, F. (2005). Automatic prominence identification and prosodic typology. *Proceedings of Interspeech*, 2005, 1813–1816.

Terken, J. (1991). Fundamental frequency and perceived prominence of accented syllables. *The Journal of the Acoustical Society of America*, *89*(4), 1768-1776.

Turk, A. and Sawusch, J. (1996). The processing of duration and intensity cues to prominence. *The Journal of the Acoustical Society of America,* 99, 3782-3790.

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, 27(2), 171-206.

Turnbull, R. (2017). The role of predictability in intonational variability. *Language and Speech*, *60*(1), 123-153.

Watson, D. G., Arnold, J. E., & Tanenhaus, M. K. (2008). Tic tac toe: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106, 1548–1557.

Wright, R. (2004). Factors of lexical competition in vowel articulation. *Papers in laboratory phonology* VI, 75-87.

Cox, J. & Henrichsen, L. (2018). An ESL teacher's guide to pronunciation teaching using online resources. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 230-238). Ames, IA: Iowa State University.

AN ESL TEACHERS' GUIDE TO PRONUNCIATION TEACHING USING ONLINE RESOURCES

Jenelle Cox, Brigham Young University

Lynn Henrichsen, Brigham Young University

ESL students view the teaching of pronunciation as an important part of their English language education, yet research has shown it is granted relatively little attention in the ESL classroom. Although many ESL teachers agree with students on the importance of pronunciation teaching, they often do not feel qualified to teach pronunciation due to insufficient training in this area. Unfortunately, time constraints and other obstacles often keep practicing ESL teachers from obtaining the pronunciation-instruction preparation they need to fulfill students' needs. These factors have led to the development of an online guide (at www.englishpronunciationguide.weebly.com) to ESL pronunciation-teaching resources. This website is designed to help teachers with inadequate pronunciation training quickly access existing pronunciation-teaching videos online. Teachers can then use these resources to educate themselves and teach specific aspects of English pronunciation to their students with greater competence and increased confidence.

INTRODUCTION

Pronunciation plays a key role in ESL learners' successful communication in English and their perceived ability in doing so (Bakar & Abdullah, 2015). Some students who have had the opportunity to take a pronunciation course have described benefits that included enhanced awareness of their own pronunciation weaknesses, increased confidence, refined listening skills, and personal tools to improve their pronunciation throughout their lives (Derwing, Munro, & Wiebe, 1998; Henrichsen & Stephens, 2015). However, too few students receive pronunciation instruction, let alone pronunciation courses, as part of their English language training. Consequently, ESL learners without specific pronunciation training may suffer from the consequences of poor English pronunciation. For example, the strong non-native accents of ESL learners often cause undesirable social interactions that include discrimination in employment and even harassment (Franklin, 2016). For this reason, ESL students typically view pronunciation as being very important and a priority (Willing, 1988) in their language education. Many ESL students continue to struggle with English pronunciation even after studying and learning the language for years (Gilakjani, 2011). This unfortunate outcome could be the result of the low priority given to explicit pronunciation teaching (Algahazo, 2015) in many ESL classes or programs. Even with students expressing their desires for pronunciation instruction, it is often included only as a minor component in speaking classes (Munro & Derwing, 2006). Yet, research shows that pronunciation should be considered an integral component of classroom activities (Gilakjani, 2011). In brief, the teaching of pronunciation—although important—

remains largely neglected in the field of English language teaching (Foote, Trofimovich, Collins, & Urzúa, 2013; Gilakjani, 2011; MacDonald, 2002; Munro & Derwing, 2006).

THE PROBLEM

One reason pronunciation teaching is not given adequate attention is that some teachers do not feel it is a necessary skill. A survey done by Grim & Sturm (2016) found that foreign language teachers did not view pronunciation as a fundamental skill to acquire and rated it as the least important when compared to other language skills. In contrast, *students* rated pronunciation as the third most important language skill, behind culture and grammar.

Another reason pronunciation teaching is neglected stems from teachers' feelings of inadequacy. Many ESL teachers do not feel qualified to teach pronunciation due to their lack of formal pronunciation training (Derwing, Diepenbroek, & Foote, 2012). The limitations felt by many teachers regarding pronunciation teaching have been documented in various contexts—by Fraser (2000) in Australia; Burgess and Spencer (2000) in the UK; Foote, Holtby, & Derwing (2011) in Canada; and Derwing (2008) in the USA. MacDonald (2002) cites several studies in Australia indicating that many teachers do not teach pronunciation because they do not feel competent to do so. Such teachers want to provide their students with the education they need to be efficient communicators, yet they lack the training themselves. For this reason, some teachers have expressed a desire for training in pronunciation teaching (Foote, Holtby, & Derwing, 2011) even though their time for professional development is limited.

If teachers do not provide pronunciation instruction, students are left on their own, and this can lead to incorrect assessment and misguided actions. For example, students tend to perceive segmentals, rather than suprasegmentals, as the area where they need to improve (Derwing & Rossiter, 2002). Nevertheless, since suprasegmentals have been shown to be the most beneficial in improving intelligibility and comprehensibility (Derwing & Munro, 2009), focusing on segmentals limits learners' progress toward intelligibility. Some teachers also overlook this key element in pronunciation teaching. A study of three instructors involved in 92 pronunciation episodes (covering 400 hours of instruction) found not even one episode of suprasegmental teaching. Teachers' overall lack of knowledge and confidence may account for this void (Burgess & Spencer, 2000). Segmental sounds represented by letters of the alphabet are easier to identify and address, and that may be one reason for students' and teachers' overlooking important prosodic features. Another reason is that it is easier to correct a single phoneme rather than the multiple words that many suprasegmentals involve (Foote, Trofimovich, Collins, & Urzúa, 2013).

POSSIBLE SOLUTIONS

Three possible solutions to the problems described above will be discussed here. The first solution is obvious—pronunciation pedagogy needs to become part of ESL teacher-preparation curricula. TESOL teacher-education programs need to be "firmly rooted in existing research" about pronunciation teaching and learning (Derwing & Munro, 2005, p. 392). This action will ensure that teachers-in-training will realize the need to teach pronunciation and through their

TESOL program become prepared to do so. However, improving ESL teacher-preparation programs still does not address the needs of the thousands of practicing teachers, most of whom are extremely busy with preparing and teaching their classes. For them, engaging in university-based teacher development related to pronunciation instruction is not a high priority or even a possibility.

A second possible solution is utilizing CAPT (Computer Assisted Pronunciation Teaching) programs that allow students to study pronunciation on their own. CAPT seems promising because it allows teachers to provide pronunciation lessons beyond their own instructional abilities (Levis, 2007). Nevertheless, there are weaknesses in most current CAPT programs. For one, computer language programs tend to exploit the strengths of the computer rather than address the real needs of L2 pronunciation learners (Neri, Cucchiarini, Strik, & Boves, 2002). These programs need to be rooted in research-based models of L2 pronunciation learning and teaching (Pennington, 1999). In addition, teachers need to be trained to use CAPT programs, which takes time that could be used to train them in pronunciation pedagogy. Furthermore, some teachers use CAPT ineffectively because of a lack of training in pronunciation *and* technology (Levis, 2007). Finally, research shows that L2 pronunciation learners progress most from recasts and prompts from an interlocutor, which computers cannot do (Gooch, Saito, & Lyster, 2016).

A third possible solution is to help teachers use existing pronunciation-teaching resources. In this regard, MacDonald (2002) urged,

> Promote existing materials. It is recommended that existing materials be promoted and made available to teachers.…To do this, they need resources and direction on how best this can be done. The recommendations thus made have been made with the view to overcoming teacher reluctance in the area of pronunciation teaching and encouraging teachers…to teach pronunciation confidently, effectively, and more often. (p.14)

To summarize thus far, this overview of relevant research literature shows that teachers generally want to help their students reach their pronunciation improvement goals but often lack the knowledge to feel confident in helping L2 learners with their pronunciation struggles (Foote, Holtby, & Derwing, 2011). Several potential solutions to this problem exist.

OUR SOLUTION

Following the recommendation to use existing materials (MacDonald, 2002) and knowing teachers' time constraints, we felt that a logical solution would be to utilize the multitude of ESL-pronunciation videos available online, focusing *not* on the thousands of online pronunciation videos for students to learn from but rather on videos that show *teachers* how to teach pronunciation to students.

Because of the multitude of English-pronunciation videos available online, choosing the best or most appropriate ones can be a daunting task for teachers who are not trained in ESL pronunciation instruction and have little time available to hunt for videos online. Therefore, we have developed a website titled *ESL Teacher's Guide to Pronunciation Teaching* (available at www.englishpronunciationguide.weebly.com) that not only links to videos designed for English

pronunciation teachers but also organizes these video links into simple, logical categories to make them easily accessible. In addition, we have previewed and pre-selected all the videos available through this website so that teachers coming to it will be led to only those videos that meet certain criteria.

DEVELOPMENT: CRITERIA FOR SELECTING VIDEOS

After viewing and evaluating close to 300 online pronunciation-teaching videos, we chose 67 that we found to be the most helpful, according to criteria explained here. First, we looked for videos designed to show teachers how to teach pronunciation to students, rather than those intended for language-learners to use on their own to improve their pronunciation. In addition, we chose videos presenting lessons on English segmentals and suprasegmentals that carry the highest functional load (Catford, 1987). We also limited our focus to those categories of English pronunciation where improvements produce the greatest gains in intelligibility (Avery & Ehlrich, 1992). Another guiding consideration was to favor videos that addressed the pronunciation difficulties experienced by the largest L1 English-learner groups—namely speakers of Chinese, Japanese, Korean, Portuguese, and Spanish. Suprasegmentals were given special attention due to their importance in ESL learners' intelligibility, but online videos designed to help teachers teach commonly mispronounced English vowels and consonants were also included. In brief, the videos we chose met as many of the following criteria as possible:

- Fit with widely recognized categories of ESL learners' pronunciation difficulties, as described in *Teaching American English Pronunciation* (Avery & Ehrlich, 1992)
- Address pronunciation features with high functional load (Catford, 1987)
- Focus on areas of difficulty associated with many ESL learners' L1s (Chinese, Japanese, Korean, Portuguese, Spanish)
- Are quick and easy for non-linguistic teachers to understand (not overly technical)
- Are interesting, even entertaining
- Feature a variety of presenters
- Provide visual support showing articulatory positions and movements (mouth, tongue, etc.)
- Use a speech rate targeted towards teachers, not L2 learners
- Provide viewer involvement—things to learn and DO, not just watch
- Run from one to ten minutes in length (although some longer exceptions were allowed because of strengths in other criterion areas)
- Give the "biggest bang for the buck" (i.e., present highly useful pedagogical skills for the targeted pronunciation feature in a short amount of time)
- Utilize a variety of approaches to teaching the desired pronunciation goals so as to fit different teachers' teaching styles and reach ESL learners with varied learning styles.

APPLICATION

The purpose of the videos that our *ESL Teacher's Guide to Pronunciation Teaching* website (at www.englishpronunciationguide.weebly.com) links users to is to give "non-pronunciation teachers" (with little or no training, experience, skill, or confidence in ESL pronunciation

instruction) the instructional models, guidance, and encouragement they need in order to gain the competence and confidence that so many ESL teachers state they lack (Yates, 2001).

Our intent is that after teachers have recognized their students' pronunciation needs, they will search our website to find online videos related to those needs. At the website, they will select, link to, and watch several videos, and then incorporate the instructional models and content into their own teaching. Ideally, teachers will feel inspired to combine techniques as appropriate and develop pronunciation lessons that help students improve their pronunciation. Put more systematically and in greater detail, we envision teachers following these steps:

1. Notice and identify their ESL student's pronunciation needs.
2. Go to the website: www.englishpronunciationguide.weebly.com (see Figure 1).
3. Choose the "Segmentals" or "Suprasegmentals" section (see Figure 2), whichever corresponds with the needs of their students.
4. Choose from three to eight videos in their pronunciation target area (see Figure 3) to view and learn from.
5. Practice and prepare their own lessons based on the models and content they have learned about in the videos they selected and viewed.
6. Implement their lessons in the classroom with creativity and confidence.
7. Try variations of the methods presented in the videos to reach students with different learning styles.

Figure 1. Home page of website.

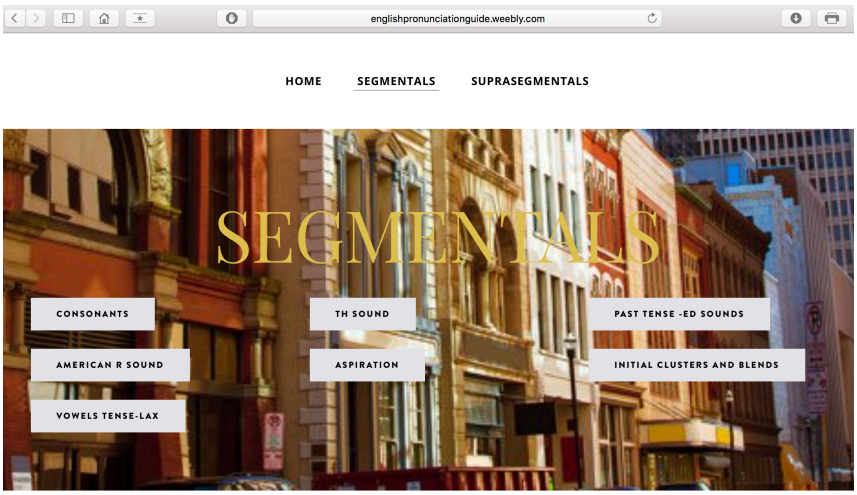


Figure 2. Segmentals directory page.

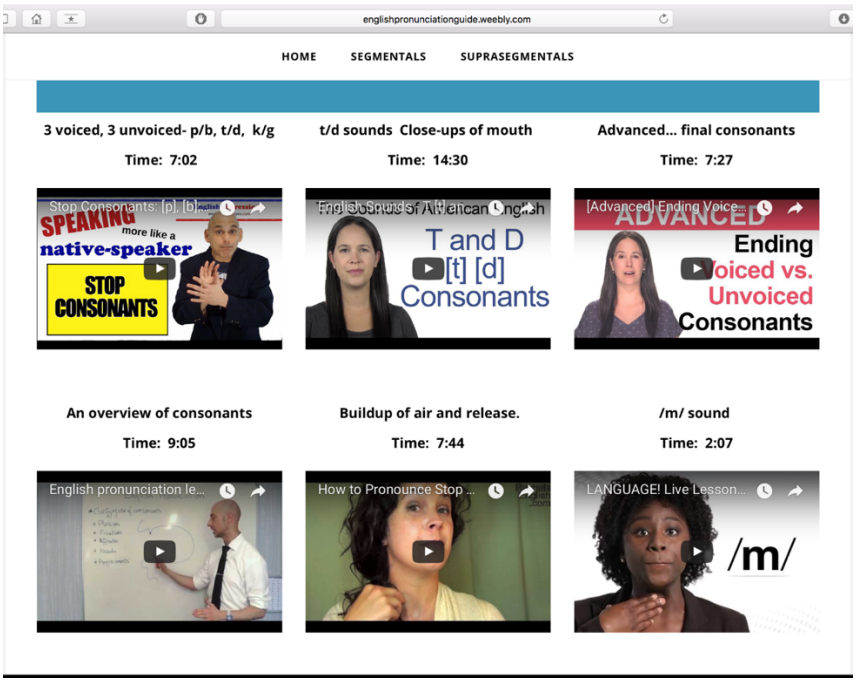


Figure 3. Segmentals video selection page.

CONCLUSION

Our hope is that teachers who use this online guide to English pronunciation teaching will feel empowered by the quality, directness, and ease of teaching high-functional-load sounds and prosody that the video models provide. We feel that giving users a reduced number of pre-selected, high-quality pronunciation-teaching videos to choose from will eliminate the overwhelming task of sorting through the countless videos online by themselves. The fact that our pre-selected videos are organized into clear, problem-oriented categories makes finding the right instructional model and correct linguistic content even easier, boosting teachers' motivation to view the pre-selected videos. Implementing what they learn from the video models will, in turn, allow "non-pronunciation teachers" to teach the needed pronunciation lessons in their classrooms with better instructional procedures, greater confidence, and increased effectiveness.

ABOUT THE AUTHORS

Jenelle Cox, a graduate of the TESOL MA program at Brigham Young University, holds a bachelor's degree in family science from BYU and a Multiple Teaching Certificate from California State University, Fullerton. Her graduate thesis addresses the need for explicit pronunciation training in TESOL preparation programs as well as pronunciation supports for practicing teachers. Jenelle has taught English in Colonia, Uruguay and in Salt Lake City, Utah. She volunteers at the Missionary Training Center, Provo, Utah helping senior couples learn Spanish. She is currently teaching Listening/ Speaking to advanced ESL students at the English Language Center in Provo, Utah. Contact information: Jenelle Cox, 6619 W. Normandy Way, Highland, Utah 84003, 714.837.6318, jenellecox@sbcglobal.net

Lynn E. Henrichsen (Ed.D., University of Hawai'i at Manoa) is a professor of TESOL in the Linguistics Department at Brigham Young University in Provo, Utah. He teaches courses in instructional methods (including the teaching of listening, speaking and pronunciation), research methods, phonetics/phonology, and ESL pronunciation. He is the author of *Pronunciation Matters* (University of Michigan Press, 1999, now available at www.pronunciationmatters.com). In his 45 years as a language teacher, he has taught Spanish and English language learners and teachers and conducted TESOL-related research in 10 different countries around the world and authored over a dozen books, 30 chapters in books, and 65 articles in professional journals. Contact information: 4040 JFSB, BYU, Provo, UT 84602, 801.422.2938, Lynn_Henrichsen@byu.edu

REFERENCES

Alghazo, S. (2015). Advanced EFL learners' beliefs about pronunciation teaching. *International Education Studies, 8*(11), 63-76.

Avery, P., & Ehrlich, S. (1992). *Teaching American English pronunciation.* New York: Oxford University Press.

Bakar, Z.A., & Abdullah, M. R. (2015). Importance of correct pronunciation in spoken English: Dimension of second language learners' perspective. *Pertanika Journal of Social Sciences and Humanities, 23*(S), 143-158.

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

Catford, J. C. (1987). Phonetics and the teaching of pronunciation*:* A systemic description of English phonology. In J. Morley (Ed.), *Current perspectives on pronunciation* (pp.83-100). Washington, DC: TESOL.

Derwing, T. M. (2008). Curriculum issues in teaching pronunciation to second language learners. In J. G. H. Edwards, & M. L. Zampini (Eds), *Phonology and second language acquisition,* 347-369. Amsterdam: John Benjamins Publishing Company.

Derwing, T. M., Diepenbroek, L. G., & Foote, J. A. (2012). How well do general-skills ESL textbooks address pronunciation? *TESL Canada Journal, 30*(1), 22-45.

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

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

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

Derwing, T. M., & Rossiter, M. (2002). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning, 13*(1), 1-17.

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

Foote, J. A., Trofimovich, P., Collins, L., & Urzúa, F. S. (2013). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal, 44*(2), 181-196.

Franklin, A., & McDaniel, L. (2016). Exploring a phonological process approach to adult pronunciation training. *American Journal of Speech-Language Pathology, 25,* 172-182.

Fraser, H. (2000). *Coordinating improvements in pronunciation teaching for adult learners of English as a second language.* Canberra, Australia: Department of Education, Training and Youth Affairs.

Gilakjani, A. P., & Ahmadi, M. R. (2011). Why is pronunciation so difficult to learn? *English Language Teaching, 4*(3), 74-83.

Gooch, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English /ɹ/ to Korean adult EFL learners. *System, 60,* 116-127.

Grim, F., & Sturm, J. (2016). Where does pronunciation stand in the 21st century foreign language classroom? Educators' and learners' views. In J. Levis, H. Le, I. Lucic, E. Simpson, & S. Vo (Eds). *Proceedings of the 7th Pronunciation in Second Language*

Learning and Teaching Conference, Dallas, TX, October, 2015, (pp. 51-59). Ames, IA: Iowa State University.

Henrichsen, L. & Stephens, C. (2016). Advanced adult ESL students' perspectives on the benefits of pronunciation instruction. In J. Levis, R. Mohammed, M. Qian, & Z. Zhou (Eds.). *Proceedings of the 6th Pronunciation in Second Language Learning and Teaching Conference. Santa Barbara, CA* (pp. 197-205). Ames, IA: Iowa State University.

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

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

Morley, J. M. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*, 481-520.

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

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning, 15*, 441-467.

Pennington, M. C. (1999). Computer-aided pronunciation pedagogy: Promise, limitations, directions. *Computer Assisted Language Learning 12*, 427-440.

Willing, K. (1988). *Learning styles in adult migrant education.* Adelaide: National Curriculum Resource Centre.

Yates, L. (2001). *Teaching pronunciation in the AMEP: Current practice and professional development.* AMEP Research Centre. Sydney, Australia: Macquarie University Press.

Elliott, N. (2018). Teaching tip: the vowel elevator: a visual-kinesthetic way to expand the vowel space. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 239-244). Ames, IA: Iowa State University.

TEACHING TIP:

THE VOWEL ELEVATOR: A VISUAL-KINESTHETIC WAY TO EXPAND THE VOWEL SPACE

Nancy C. Elliott - University of Oregon

The concept of the vowel space can be difficult to envision for ELLs, teachers -in-training, and even students of linguistics. The English vowel system is a particular challenge for learners, partly due to the precise target areas required for a system with five levels of vowel height. Teachers can help students with this challenge by using the metaphor of the Vowel Elevator, a visual-auditory-kinesthetic method of pronunciation practice that helps students comprehend and expand their vowel space using movement, sight, and sound, while keeping the whole system in the basic organization of high-mid-low and front-central-back articulation. The mouth is represented as a building with elevators that move between stories – with different languages having different numbers of stories– and learners practice the movements with simple hand gestures that accompany their tongue movements as they produce the sounds and view images of the elevator stopping at particular floors. Learners can visualize the vowel space, move around in it more accurately, and attach sound associations to the movements.

The challenge of understanding the vowel space

Language learners, teachers-in-training, and students of linguistics often find it difficult to understand the concept of the vowel space. In one introductory linguistics textbook (Fromkin, Rodman & Hyams 2003), the authors reassure the learner with the following words: "You may not understand at first what we mean by 'front,' 'back,' 'high,' and 'low' vowels, but we encourage you to persist. It will come" (p. 252). In comparison, the characteristics that distinguish most consonant sounds – voicing, place and manner of articulation - are more straightforward to describe, and suprasegmentals are often defined by simple binary parameters, e.g., stressed-unstressed; higher pitch-lower pitch; longer vowel-shorter vowel. The English vowel system in particular is also a challenge for learners due to the precise target areas required for a system with five levels of vowel height.

Conceptualizing tools for teaching the vowel space

Various conceptualizing tools exist for teaching or describing the vowel space, including the vowel triangle, haptic clockface movements, and Color Vowel® approach. The vowel triangle or quadrilateral, commonly presented in introductory linguistics and foreign language textbooks, is a visual mapping of vowel symbols on a two-dimensional space that represents either the oral cavity or acoustic correlates. Approaching the vowel space concept quite differently, the creators of haptic-integrated pronunciation instruction use the clock face metaphor, whereby learners move their arms to the positions of analog numbers correlated with particular vowel sounds as they say words containing those sounds (Acton, Baker, Teaman & Burri 2012). The

Color Vowel® approach offers teachers and learners a visual-aural method of associating vowel sounds with assonant words for particular colors (Taylor & Thompson 1999, 2015).

I would like to contribute to this small kit of tools the concept of the Vowel Elevator, which uses metaphor, movement, and imagery to help learners visualize the vowel space, move around in it more accurately, and attach sound associations to the movements. In this metaphor, the mouth is presented as a building with elevators that move between stories, with different languages having different numbers of stories. Learners view images of the elevator stopping at particular floors and practice the movements with simple hand gestures that accompany their tongue movements as they produce the sounds. Thus, along with the two senses usually employed in learning to make sounds (hearing and proprioception, or the sense of body position), the multisensory channels of vision, kinesthesia, and touch are added in order to enhance learning by strengthening memory formation (Shams & Seitz 2008).

The Vowel Elevator in practice: the American English Vowel System

To introduce the vowel space in an ESL classroom, students are given an illustrated slide show that presents the following ideas:

1. *(Image of skyline with skyscrapers in different shapes)* Languages are different buildings.
2. *(Image of hand pushing button on 5-floor elevator keypad)* The English vowel building has five stories.
3. *(Image of three-floor elevator keypad)* Other languages have fewer.
4. *(Image of three-story house aligned with image of Spanish vowel triangle)* Japanese and Spanish language buildings have three floors.
5. *(Image of tall two-story house aligned with image of Arabic vowel triangle)* The Arabic language is a two-story building.
6. *(Image of five-story house aligned with image of English vowel quadrilateral)* The English vowel building has five floors.
7. *(Image of a building with three elevators)* English has a front elevator, a back elevator, and a service elevator in the middle.
8. *(Image of hallway with open elevator door)* Let's get on the front elevator and ride to the top.

At that point in the slide show, students are introduced to the continuum of American English front vowels and back vowels by 'going up and down the elevators.' While students view a photo of elevator buttons with the up arrow illuminated, they are directed to say "ah----ee," in order to become aware of the tongue movement from the 'bottom floor' to the 'top floor.' Then they are told to move their hand from a low position to a higher one as they articulate "ah----ee" once again. As they are shown a photo of elevator buttons with the down arrow illuminated, they move their hand back down and say "ee----ah."

Continuing the metaphor, students then 'stop at some floors on the way to the bottom' and are introduced to the five front vowels, as shown in the following slides:

The front elevator has these five vowels:

5th floor / i / as in "key, beat"
4th floor / ɪ / as in "him, bit"
3rd floor /ev/ as in "okay, bait"
2nd floor / ɛ / as in "them, bet"
1st floor / æ / as in "hand, bat"

Listen to these five words:
beat – bit – bait – bet – bat

Did you hear the vowel elevator stop at each floor?

Let's take the English elevator down one floor from the top.

In English, the fifth floor is / i / (as in "key").
Say these words:
team
leave
sheep
heat

In English, the fourth floor is / ɪ / (as in "him").
Say these words:
Tim
live
ship
hit

Make your tongue-elevator go down one floor...

- Move your hand as we take the elevator from the fifth floor to the fourth floor. Say the vowels as we travel:
  / i / ..... / ɪ /
- Do it again: / i / ..... / ɪ /.
- One more time: / i / ..... / ɪ /

Make your tongue-elevator go down one floor...

Elevator Game

Find a partner. Say one of the words below. Can your partner guess which floor you're on?


• team	Tim
• sheep	ship
• leave	live
• heat	hit

Let's take the English elevator from the 4th floor to the 2nd floor.

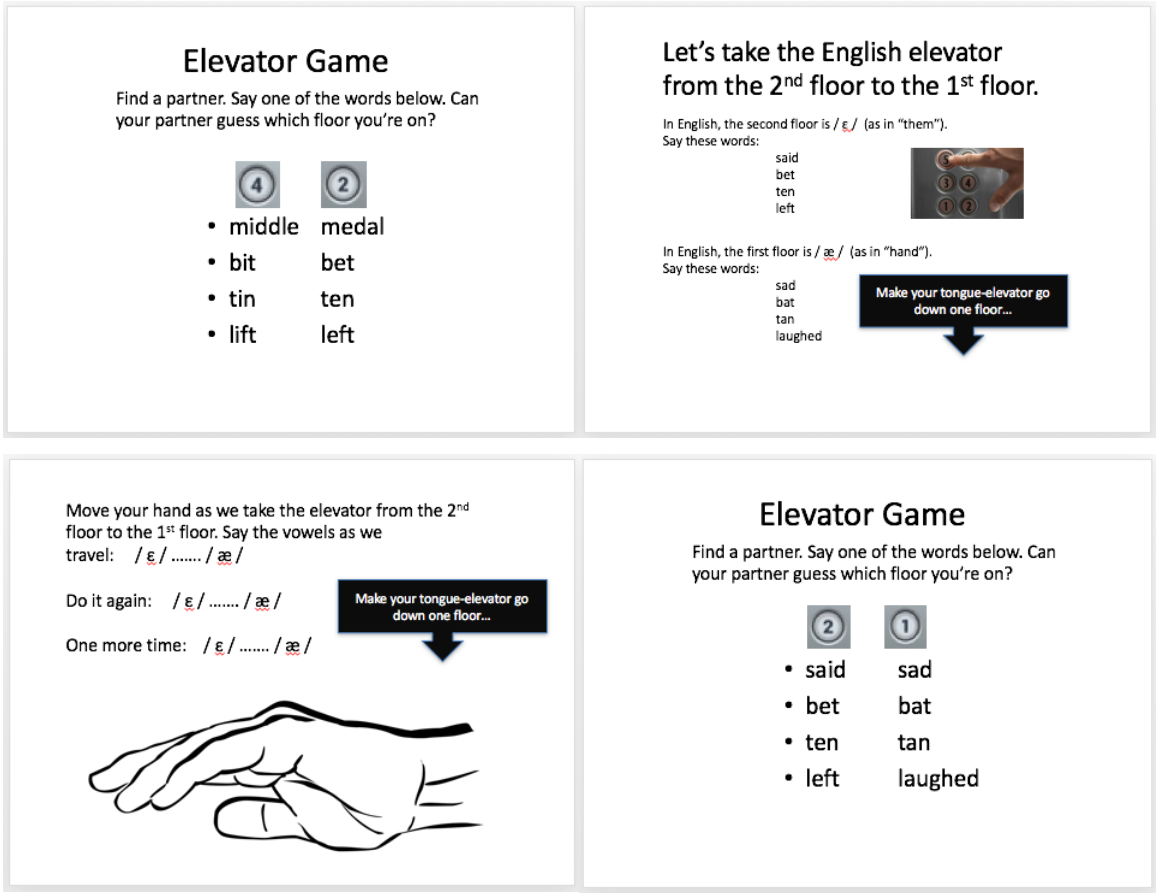
In English, the fourth floor is / ɪ / (as in "him").
Say these words:
bid
disk
middle
lift

In English, the second floor is / ɛ / (as in "them").
Say these words:
bed
desk
medal
left

Make your tongue-elevator go down two floors...

Find the sound / ɪ / in your mouth (think "him").
Now, move your hand as we take the elevator from the 4th floor to the 2nd floor.
Say the vowels as we travel:
/ ɪ / ..... / ɛ /
Do it again: / ɪ / ..... / ɛ /

Make your tongue-elevator go down two floors...



The example slides above are used to demonstrate and practice front vowels, and include vertical tongue movement, accompanying hand movement, and 'elevator games' for pair or group practice. A similar sequence of slides is shown in a subsequent class period for the back vowels. During this sequence, it is pointed out that when taking the back elevator, most people in the U.S. Midwest and West skip the 2nd floor and land on the 1st floor instead. This is a convenient way to explain the *cot-caught* merger, in other words the regional falling-together of low-back /*a*/ and mid-back /ɔ/.

When it comes to English schwa, the elevator metaphor extends to include a central, express elevator that goes directly to a very popular floor that is 'right in the middle of everything.' Students can then play an elevator game in which they identify which elevator the speaker is on:

When you enter the English building,
you have to choose the right elevator.

- /æ/ is the front elevator.
- /ə / is the central elevator.
- /ɑ/ is the back elevator.

Front, central, or back?

- Pick a partner. Say a word. Can your partner guess which elevator you are on?

Front	Central	Back
lack	luck	lock
Dan	done	Don
cap	cup	cop
bag	bug	bog
hat	hut	hot

Another feature of English vowels that can be illustrated and practiced using the elevator metaphor is the diphthongization of mid-front [ey] and mid-back [ow]; the off-glide is compared to the jiggle of an elevator that always bounces on the third floor:

The elevators bounce a little on the third floor.

Listen to the front elevator: /ey/
Listen to the back elevator: /ow/

Can you hear the bounce?

Now try it:
Say /ey/ and move your hand on the bounce.
Say /ow/ and move your hand on the bounce.

Say these third-floor words.
Make sure the elevator bounces!
Move your hand as the sound bounces upward.

FRONT ELEVATOR	BACK ELEVATOR
gray	snow
place	rose
eight	boat
remain	phone

Final thoughts: extended metaphors as teaching tools

Metaphors are common in language description and teaching. For illustrating verb tenses, time is a line; in phonetics, some consonants are glides or liquids; pitch is high or low; intonation rises and falls; in writing, a paragraph is often described as a sandwich. The vowel space is real, however. Vowel space descriptions - high, mid, low, front, central, back - are literal, not metaphorical, but it is a challenge for learners that these parameters basically cannot be seen and are not easy to sense with proprioception. No metaphor is a perfect representation of reality, but the Vowel Elevator metaphor maps well onto reality: The oral cavity is a building; the tongue is a series of elevators; the building's stories are vowel heights. English has a front elevator, a back elevator, and an express elevator in the middle that goes directly to the very important and popular "schwa" area of the building. The front and back elevators bounce a little when they stop on the third floor (a reference to diphthongized mid vowels), and when taking the back elevator, most people in the U.S. Midwest and West skip the 2nd floor, and head to the 1st floor instead (a reference to the *cot-caught* merger).

Taking this metaphor and running with it, learners can compare the L2 vowel building with that of their L1. They experiment with new mouth-shapes and practice arriving at various floors of

the vowel building. They work with a partner to produce and identify vowel sounds with the activity "*Guess which floor I'm on*." Students respond to the combination of movement, sound, and visual imagery, which helps keep them engaged and piques learners' curiosity about the sound system of the language that they are learning.

ABOUT THE AUTHOR

Nancy Elliott teaches ESL at the University of Oregon, and has also taught languages and/or linguistics at the University of Kansas, Universität Heidelberg, Indiana University, and Southern Oregon University. She has a Ph.D. in linguistics from Indiana University, where she specialized in English dialectology, sociolinguistics, and the history of English pronunciation. Her research interests include English rhoticity, accent in the media, and listener judgements of comprehensibility and intelligibility.

American English Institute
1787 Agate Street
5212 University of Oregon
Eugene, OR 97403-5212
(541) 346-3945
nce@uoregon.edu

REFERENCES

Acton, W., Baker, A., Teaman, B., & Burri, M. (2012, August). *Preliminaries to haptic-integrated pronunciation instruction*. Paper presented at the 4th Annual Pronunciation in Second Language Learning and Teaching Conference, Vancouver, BC.

Fromkin, V., Rodman, R., & Hyams, N. (2003). *An introduction to language*. Boston: Thomas and Heinle.

Shams, L., & Seitz, A. (2008). *Benefits of multisensory learning*. Trends in Cognitive Science 12: 411-417.

Taylor, K., and Thompson, S. (1999, 2015). *The Color Vowel Chart*. Santa Fe NM: English Language Training Solutions. https://www.colorvowelchart.org.

McGregor, A. (2018). Teaching tip: The tic tac trick to teach the American English articulatory setting. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 245-251). Ames, IA: Iowa State University.

TEACHING TIP

THE TIC TAC TRICK TO TEACH THE AMERICAN ENGLISH ARTICULATORY SETTING

Alison McGregor, Princeton University

INTRODUCTION

Second language pronunciation instruction typically begins with a focus on articulation or the manner in which to produce individual vowels and consonants. Lacking from the teaching and learning of English pronunciation, however, is instruction on the language's *articulatory setting* (AS), or default position for articulators. Surprisingly, since as early as Sweet (1890), it has been pointed out that unique to each language are its underlying tendencies, habitual configurations, or default positioning of the articulators (lips, jaw, tongue, and velum). The concept of language-specific underlying tendencies was termed articulatory setting by Honikman (1964) and defined as the "gross oral posture and mechanics both external and internal, requisite as a framework for the comfortable, economic, and fluent merging and integrating of the isolated sounds…" (p. 73). Consequently, AS offers increased efficiency and effectiveness of pronunciation instruction because it underlies sound production and provides the tools to over-come cross-linguistic transfer tendencies, which often impede L2 pronunciation.

In addition, according to the Noticing Hypothesis (Schmidt, 1990), conscious awareness is necessary to convert input to intake. Applied to pronunciation instruction, this means awareness is needed to begin the learning process in converting consciousness about production mechanisms to actual fluent production. Specifically, learners need both the *awareness of noticing*, that is, the "conscious registration of the occurrence of some event" and *awareness as understanding*, "recognition of a general principle, rule or pattern" (Schmidt, 1995, p 29). The Tic Tac trick provides a practical technique for introducing the noticing and understanding of AS for American English. The aims of this teaching tip are to: (a) introduce articulatory settings and the settings specific to American English and (b) describe the use of a Tic Tac to help students become conscious of the American English settings.

Background

According to Sweet (1890), "Every language has certain tendencies which control its organic movements and positions, constituting its organic basis or the basis of articulation. A knowledge of the organic basis is a great help in acquiring the pronunciation of a language" (p. 69). In addition to *organic movements and positions*, the concept of AS has been referred to as organic base (Wittig, 1956), Artikulationsbasis (Arnold & Hansen, 1975), voice quality settings (Esling & Wong, 1983), postural configurations (Gick, Wilson, Koch, & Cook, 2004), and articulatory posture (Roach, 2014)[1]. In spite of the variation in terminology, there is consistent historical and contemporary recognition that each language has a unique set of underlying mechanisms, or default settings, which ultimately enable language-specific speech. There is also evidence that underlying AS is a language's neutral vowel, which is produced with the articulators in a configuration closest to the AS of that language (Gick, et al., 2004). Even postures (settings) during grammatical pauses have been found to have mechanical advantage to facilitate efficient

[1] See Jenner (2001) for a more thorough historical description of the concept.

postural motor control of articulators (Ramanarayanan, Lammert, Goldstein, & Narayanan, 2014). This information suggests that AS underlies the positions of 'schwa' and typical filled pauses of "uh" and "um" in American English and the position plays a significant role in efficient speech production. Recently, actor Christopher Aruffo gave a *TEDxMcMasterU* talk describing his experience learning different accents by watching and mimicking. Aruffo (2012) skillfully demonstrates a direct connection between muscle movements and settings in relation to their implications for changes in how one's speech sounds.

In sum, AS can be viewed as the "default settings" for a language and offers pronunciation instruction a corner stone for efficient learning and effective production of a language. Since AS can be seen as a foundation of pronunciation, there is a strong rationale to include it in pronunciation instruction. The underlying theoretical basis of this teaching tip that will be described below is noticing, that is, students become aware of and understand what to do for new motor-skill development. The pedagogical premise for this teaching tip embraces effective pronunciation instruction as a systematic and scaffolded[2] learning process that includes awareness-raising (discovery), introduction of explicit information (declarative knowledge), guided practice, self-monitoring and assessment, and feedback. As a teaching technique, the Tic Tac trick can be used for discovery and facilitation of these key components in the learning process.

What Teachers Need to Know about Articulatory Setting

Introduction to AS and the American English Settings. When two languages differ, attempting to master the pronunciation of one while maintaining the articulatory setting of the other will impede production accuracy (Honikman, 1964). Articulators include the lips, the jaw, the tongue, and the velum, or soft palate, which is located at the back of the roof of the mouth and can be referred to simplistically as the *throat*. These articulators can be divided into *external settings* (the lips and the jaw) and *internal settings (*the tongue and the velum). Table 1 shows the location of the articulators, the articulators, the potential settings for each articulator, and the settings for American English, respectively. The general AS descriptions for American English include a) relaxed lips (not protruding or rounded), b) relaxed and loosely closed jaw position, c) relaxed neck and throat, and d) centered tongue, slightly forward.

Since AS is language-specific, learners' articulatory settings will naturally be influenced by the settings used in their L1. Consequently, AS-related issues (a) cause L1 cross-linguistic transfer, (b) impede effective production accuracy of vowels and consonants, and (c) frustrate learners who are completely unaware of its influence –including what to change and how to change it. Given these challenges, the rationale for the teaching and learning of AS includes its potential to reduce overall L1 phonetic transfer, increase effectiveness and efficiency in the process of learning L2 sounds, and minimize learner frustration.

Table 1

Introduction of Articulatory Settings

[2] See Sardegna & McGregor (2013) on *Scaffolding students' self-regulated efforts for effective pronunciation practice.*

Location	Articulators	Potential Settings	English AS
External	Lips	protruding/spread; rounded/unrounded	relaxed (i.e., not protruding, not spread, not rounded)
	Jaw	open/closed; tense/relaxed	(Loosely) closed; relaxed
Internal	Tongue	height-high/mid/low; forwardness/backness; tip/blade-tense/lax;	mid/central; a bit forward; lax;
	Throat (velum)	tense/relaxed; closed/open	relaxed; open

What to expect and how to assess. To use AS in pronunciation instruction, a comparison must be made between the settings of the languages involved in order to identify the relative differences in the settings of the articulators. For example, Danish learners of English should be advised that in English, there is less lip activity and a more relaxed jaw than in Danish, and the tongue tip/blade is a bit more tense (Collins & Mees, 1998). Descriptions for learners of different L1s will naturally differ. For example, a native Mandarin speaker learning English will need different setting adjustments from those of a native French speaker learning English. Teachers can expect some common articulator tendencies for the same L1s although there might be slight individual differences.

Teachers need to look for the internal and external setting of their L2 learners' production to identify the L1's AS influence. Obviously, the external settings of the lips and jaw are easier to observe directly and describe. In general, the questions for teachers to keep in mind about AS are: What are the articulators doing and how does that compare with what they are supposed to be doing? To begin observing AS and assessing the challenges it imposes, teachers can keep the following issues in mind. First, look at the students' general lip position: Are the lips protruding, spread, or rounded? Next, consider the jaw position: How open or closed is the jaw, and does it seem tense or relaxed? For the internal settings, the teacher will have to develop an ear for the influences of tongue and throat positions. To identify tongue-setting accuracy, this is like listening for tongue positions to identify vowel accuracy. For example, the position difference between /a/ and /ə/ relates to tongue height (and possibly backness/forwardness). The tongue setting includes tongue parts (tip/blade, middle, back (root)), as well as positions and tenseness[3]. The velum presents a similar challenge. The instructor will have to consider the neck/throat area. Does the area seem tense, and do sounds seem to come from the back, middle, or front of the mouth? Do sounds seem to come from the throat? Do they sound hollow? Answers to these questions will indicate settings related to the velum (neck/throat area).

In summary, languages differ in their articulatory settings--that is, "the overall manner in which the speech organs are held which underlies articulatory movements superimposed on them" (Collins & Mees, 1998, p. 415). Understanding AS will help teachers provide fundamental instruction and trouble-shoot challenging pronunciation difficulties in which the AS of the L1

[3] See Messum & Young (2017) article titled *Bringing the English Articulatory Setting into the classroom: (1) the tongue* for more thorough description of the tongue.

influences the L2 production accuracy.

How to Facilitate AS training in the Classroom

Preliminaries. Before introducing AS, teachers are encouraged to consider the following.

Description:	This activity draws students' attention to the external and internal articulators of American English.
Tip:	Use a Tic Tac to shift students to sensing-mode; create an experience for seeing and feeling the difference between students' L1 and English AS.
Level:	Articulatory settings will be most appropriate for low-intermediate to advanced level learners.
Materials:	Tic Tacs, cell phone cameras (in selfie-mode) or mirrors, a checklist or comparison table (See Tables 2 & 3).
Prerequisites:	Students will need to be familiar with the following vocabulary: lips, jaw, tongue, throat (visual aids are recommended, especially for lower-level students); notice, watch, feel; tense, relaxed, open/closed, forward/back, middle/center, open/closed.

Introducing AS in the Classroom. The concept of articulatory settings can be introduced in a number of ways. For example, the teacher can use a computer analogy, explaining that every language has *"default settings."* For lower-level students, the teacher might draw the foundation or frame of a house or use a sports analogy of the form in shooting a basketball. Alternatively,

Part 1: Directions to Students

1. Everyone take a Tic Tac. (Don't chew it!)
2. Prepare to look in your cell phone camera (or use a mirror) to watch your production.
3. Put the Tic Tac in your mouth on the middle of your tongue.
4. Slowly open your mouth and say "uh" (/ə/).

Part 2: Directions for Teacher

5. Listen to students' production of the schwa and advise on accuracy. Tongue height and forwardness/backwardness are common problems. Lip rounding or jaw opening or width may need
6. Next, use Table 3 noticing question prompts to direct students' attention to articulators.
7. Ask students to share what they notice (see and feel) is similar or different about the settings for each articulator (lip, jaw, tongue, throat).
8. Based on observations, make recommendations on what adjustments would improve students' production.

teachers can start by asking students what Americans say when they pause or hesitate, (Uh/Umm…) and contrast that with what the students say in their first language when they pause or hesitate to speak. Making a comparison of these neutral vowels can begin raising awareness that a fundamental difference (even if subtle) does indeed exist.

Instruction for the Tic Tac Trick. The initial learning objectives of the Tic Tac trick are to raise student awareness of articulatory settings and allow students to discover the American English articulatory settings.

For the Tic Tac trick to be most effective, the instructor will need to assist students by assessing the settings of the articulators and then convey feedback as needed in terms of what to change. In the next section, a comparison activity is described.

Table 2

Noticing Questions for Tic Tac Trick

✔	**Articulators**	**Noticing Questions**	**American English Settings**	**What do I need to change?**
	Lips	What do you notice about the lips? Are they sticking out/rounded?	Relaxed	
	Jaw	Is the jaw tense or relaxed? Does it seem more open or closed in the back?	closed; relaxed	
	Tongue	Where is your tongue-is it high/low; forward/back; tense or relaxed?	mid/central; a bit forward; tip - relaxed	
	Throat	Is your throat relaxed or tense?	relaxed; open	

Comparison & Contrast of English with Students' L1 Language(s). It is extremely important to keep in mind that AS is actually a relative concept. In other words, students will benefit most when they recognize the fundamental similarities and/or difference in the positions of the articulators between English and their native language. To highlight the fact that AS exists or to create a simple comparative assessment task, ask students to video record (cell phone in selfie-mode) themselves saying a few words or short phrases in their first language and then saying the English equivalent. Next, ask them to play and compare the two recordings with sound on and off to notice the similarities and differences in their articulators. The teacher will want to follow the advice in the above section on what to expect and how to assess. If a student's L1 and English AS appears similar, the teacher will then need to identify whether the student is simply using L1 settings to produce English or if the settings are, in fact, similar. Table 3 provides a comparison guide for teachers and students to use.

Table 3

Comparison Chart for Students L1 versus American English AS

Location	Articulators	Potential Settings	Observations for L1/other language	Observations for English
External	Lips	Protruding/spread; rounded/unrounded		relaxed
	Jaw	Open/closed; tense/relax		Closed; relaxed
Internal	Tongue	Height-high/mid/low; flat/hollowed/narrow; Forward/Back; Tip/blade-tense/lax		mid/central; a bit forward
	Throat	Tense/relax; closed/open		Relaxed; open

CONCLUDING REMARKS

Given the long history of recognizing language-specific settings and the potential in providing fundamental articulatory positions, AS offers teachers and students a secret to more effective and efficient pronunciation teaching and learning. With easy access to video recording, external settings can easily be captured and identified. For internal settings, the Tic Tac trick offers a quick, easy, and inexpensive technique to draw learners into a sensing mode to raise awareness of the tongue's many potential positions. The Tic Tac trick technique is offered to teachers in the hope that they will add it to their pronunciation teaching toolbox and improve the efficiency of pronunciation training.

ABOUT THE AUTHOR

Alison McGregor is an instructor for the International Office at the University of Texas at Austin. She teaches ITA and graduate level oral communication skills courses. In addition, she conducts pronunciation teacher training courses and works as a communication consultant specializing in accent modification for the McCombs School of Business MBA program. Her research focuses on factors that promote effective pronunciation instruction, and American English intonation.

Alison McGregor, University of Texas at Austin
1611 W 39th 1/2, Apt 107, Austin, TX 7875
(512) 779-3748
mcalison@utexas.edu

REFERENCES

Arnold, R., & Hansen, K. (1975). Englische Phonetik, *Verlag Enzyklopadie*. Leipzig.

Aruffo, C. (2012). *TEDxMcMasterU - Christopher Aruffo* [Video file]. Retrieved from: http://www.youtube.com/watch?v=TDxFrwkiHIw

Collins, B., & Mees, I. (1995). Approaches to articulatory setting in foreign-language teaching. *Studies in General and English Phonetics*, 415-424.

Esling, J. H., & Wong, R. F. (1983). Voice quality settings and the teaching of pronunciation. *Tesol Quarterly, 17*(1), 89-95.

Honikman, B. (1964). Articulatory settings. In D. Abercrombie, D. B. Fry, A. D. MacCarthy, N. C. Scott, & J. L. Trim (Eds.), *In Honour of Daniel Jones* (pp. 73-84). London: Longman.

Gick, B., Wilson, I., Koch, K., & Cook, C. (2004). Language-specific articulatory settings: evidence from inter-utterance rest position. *Phonetica, 61*(4), 220-233.

Jenner, B. (2001). 'Articulatory Setting': Genealogies of an idea. *Historiographia Linguistica*, *28*, 121-141.

McGregor, A. & Reed, M. (2018). Integrating pronunciation into the English language curriculum: A framework for teachers. *CATESOL Journal, 30*(1), 69-94.

Messum, P., & Young, R. (2017). Bringing the English articulatory setting into the classroom: (1) the tongue. *Speak Out!*(57), 29-53.

Ramanarayanan, V., Lammert, A., Goldstein, L., & Narayanan, S. (2014). Are articulatory settings mechanically advantageous for speech motor control. *PLOS One 9*(8). Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104168

Sardegna, V. & McGregor., A. (2013). *Scaffolding students' self-regulated efforts for effective pronunciation practice.* Paper presented at the Pronunciation in Second Language Learning and Teaching Conference, Ames, IA: Iowa State University.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158. doi: https://doi.org/10.1093/applin/11.2.129.

Schmidt, R. W. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning *Attention and Awareness in Foreign Language Learning* (pp. 1-63). Honolulu, Hawaii: University of Hawaii, Second Language Teaching & Curriculum Center.

Sweet, H. (1890). *A primer of phonetics* (2nd ed.). Oxford: Clarendon Press.

Wittig, K. (1956). *Phonetik des Amerikanischen Englisch*. Heidelberg: Carl Winter.

Holt, E. (2018). Teaching tip: The use of MRI and ultrasound technology in teaching about Spanish (and general) phonetics and pronunciation. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 252-262). Ames, IA: Iowa State University.

TEACHING TIP:

THE USE OF MRI AND ULTRASOUND TECHNOLOGY IN TEACHING ABOUT SPANISH (AND GENERAL) PHONETICS AND PRONUNCIATION

D. Eric Holt, University of South Carolina

To teach pronunciation, it is useful for students to learn about articulatory phonetics, and to practice manipulating features of consonants and vowels. While diagrams are often presented to achieve this, these are static, and it is helpful to visualize movements of articulators via useful companions such as websites that involve functional MRI and ultrasound technology. One limitation of these is that they only present someone else's production. While fMRI is not currently a practical tool for classroom use, ultrasound technology may aid the student during the production of many potentially problematic speech sounds. In this teaching tip, an overview of online fMRI and ultrasound tools is given, with indications of their practical application.

INTRODUCTION

To teach pronunciation, it is useful for students to learn about articulatory phonetics, and this usually occurs in dedicated courses where pronunciation is embedded within the slightly larger context of phonetics and phonology. Students in these courses are instructed in the basics of the International Phonetics Association chart and symbols for transcription for both consonants (place and manner of articulation, voicing) and vowels (tongue height, backness, lip rounding, tenseness), and they usually practice identifying and manipulating these features. Often, diagrams are provided or presented to visualize these abstract characteristics. However, these are static, and being able to visualize movements of articulators can be very helpful to proprioception and development. To this end, useful companions that involve functional MRI and ultrasound technology may be employed, such as the websites "real-time MRI IPA charts", "eNunciate!" and "Seeing Speech: IPA Charts."

While these real-time anatomically-explicit views can be highly illuminating, they have the significant limitation that they only present someone else's production. fMRI is not currently a practical tool for classroom or language lab use, but ultrasound technology is lightweight, portable and relatively affordable, and can be used in the classroom in either group or individual settings. Students can directly see the position of their own tongue during the production of many speech sounds, and the real-time ultrasound images allow them to visualize themselves articulating sounds in new places and aid them to avoid pitfalls of tense/lax in vowels, clear vs. dark /l/, proper placement of palatals, and problematic sequences of sounds.

In this teaching tip, an overview of online fMRI and US tools is given, with indications of their practical application. On site, participants were given a live demonstration of ultrasound and afforded practice using the Interson 7.5MHz USB-laptop-connected wand.

Figure 1. SeeMore Interson 7.5MHz USB-laptop-connected wand.

SEGMENTS HELPFUL TO VIEW WITH IMAGING

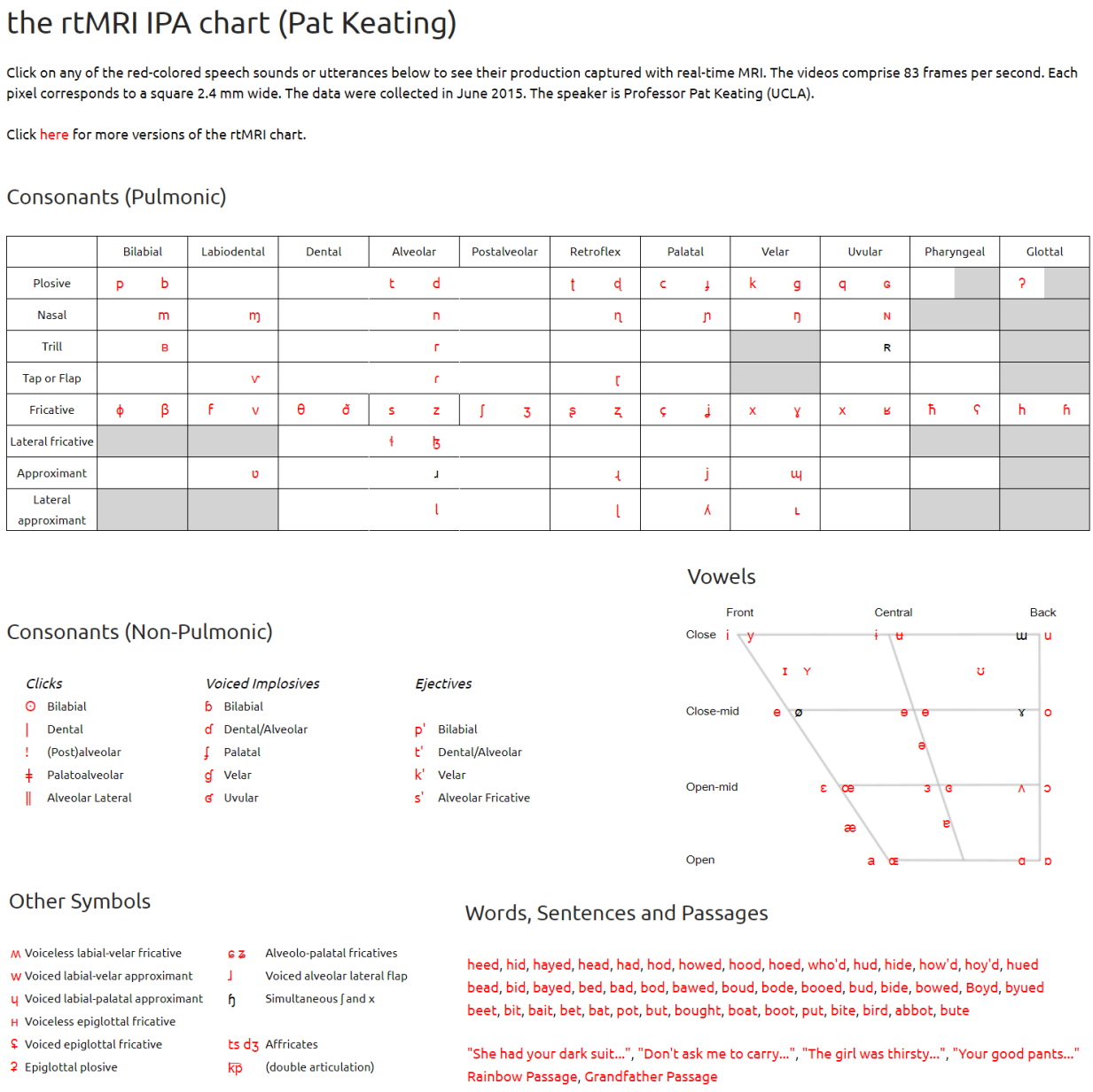
Below is a non-exhaustive list of segments, segment types and sequences for which it is especially useful to learners to visualize the production of speech, either via recorded fMRI or ultrasound, and to then visualize their own production using ultrasound. During the actual conference session, a sampling of these was demonstrated and practiced at the conclusion of the tip.

- tense vs. lax vowels: [ɛ] vs. [e]; [ɔ] vs. [o]; [ɪ] vs. [i]; [ʊ] vs. [u]
- diphthongized vs. simple vowels (diphthongs vs. monophthongs)
- tap vs. trill "r" (with M-mode) vs. bunched or retroflex "r" ([ɹ])
- *r*-colored vowels
- clear/alveolar [l] vs. dark/velarized [ɬ]
- [n] vs. [l] (cross-sectional view)
- Palatal segments vs. sequences of alveolar + palatal glide:
    - [ɲ] vs. [n+j] (*cañón* vs. *canyon*) (one gesture)
    - [ʎ] vs. [l+j] (*millón* vs. *million*) (one gesture)
        - also *el yate, un llorón* (where in Spanish the lateral or nasal anticipates the place of articulation of following palatal sound)
- alveolar vs. dental [t, d]
- Stops/occlusive/plosives vs. fricatives/spirants/approximants, e.g.,
    - [g] vs. [ɣ] (*gol* vs. *su gol*)
    - [k] vs. [x] (*acá* vs. *ajá*)
- sibilants ([s] with tongue tip up or down; cross-sectional view)

SAMPLE RESOURCES AND TOOLS

fMRI

The "span | speech production and articulation knowledge group" at the University of Southern California hosts the site at http://sail.usc.edu/span/rtmri_ipa/index.html. Here, there are a real-time MRI IPA charts with clickable IPA symbols that call up videos of both consonants and vowels, and by a variety of male and female speakers.



the rtMRI IPA chart (Pat Keating)

Click on any of the red-colored speech sounds or utterances below to see their production captured with real-time MRI. The videos comprise 83 frames per second. Each pixel corresponds to a square 2.4 mm wide. The data were collected in June 2015. The speaker is Professor Pat Keating (UCLA).

Click here for more versions of the rtMRI chart.

Consonants (Pulmonic)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k g	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Consonants (Non-Pulmonic)

Clicks		Voiced Implosives		Ejectives	
ʘ	Bilabial	ɓ	Bilabial	pʼ	Bilabial
ǀ	Dental	ɗ	Dental/Alveolar	tʼ	Dental/Alveolar
ǃ	(Post)alveolar	ʄ	Palatal	kʼ	Velar
ǂ	Palatoalveolar	ɠ	Velar	sʼ	Alveolar Fricative
ǁ	Alveolar Lateral	ʛ	Uvular		

Vowels

Other Symbols

ʍ	Voiceless labial-velar fricative	ɕ ʑ	Alveolo-palatal fricatives
w	Voiced labial-velar approximant	ɺ	Voiced alveolar lateral flap
ɥ	Voiced labial-palatal approximant	ɧ	Simultaneous ʃ and x
ʜ	Voiceless epiglottal fricative		
ʢ	Voiced epiglottal fricative	ʦ ʤ	Affricates
ʡ	Epiglottal plosive	k͡p	(double articulation)

Words, Sentences and Passages

heed, hid, hayed, head, had, hod, howed, hood, hoed, who'd, hud, hide, how'd, hoy'd, hued
bead, bid, bayed, bed, bad, bod, bawed, boud, bode, booed, bud, bide, bowed, Boyd, byued
beet, bit, bait, bet, bat, pot, but, bought, boat, boot, put, bite, bird, abbot, bute

"She had your dark suit...", "Don't ask me to carry...", "The girl was thirsty...", "Your good pants..."
Rainbow Passage, Grandfather Passage

Figure 2. "span | speech production and articulation knowledge group" sample real-time MRI IPA chart. Note: See Figure 3 for a video example.

Figure 3. Sample sentence production.

http://sail.usc.edu/span/rtmri_ipa/videos/je_2015/the_girl_was_thirsty.html

Another excellent resource is the MRI capture of the lip-synch of pop singer Adele's song *Hello* by speech researcher Nadine Lavan of the Royal Holloway Vocal Communication Laboratory, University of London: http://www.carolynmcgettigan.com/media, image capture here (and also available on Facebook here):

Figure 4. MRI capture of the lip-synch of Adele song *Hello*. Royal Holloway Vocal Communication Laboratory.

ULTRASOUND IMAGING

Two excellent resources are the websites "eNunciate!" and "Seeing Speech: IPA Charts."

eNunciate! (http://enunciate.arts.ubc.ca/) is a visual language learning tool developed at the University of British Columbia with the aim of applying ultrasound technology to test biovisual feedback in the context of the language learning (Figure 5). It includes ultrasound overlay videos, in which ultrasound images of tongue movements in speech are superimposed on videos of a face to allow users to visualize how facial and tongue muscles are coordinated. (eNunciate! Homepage)

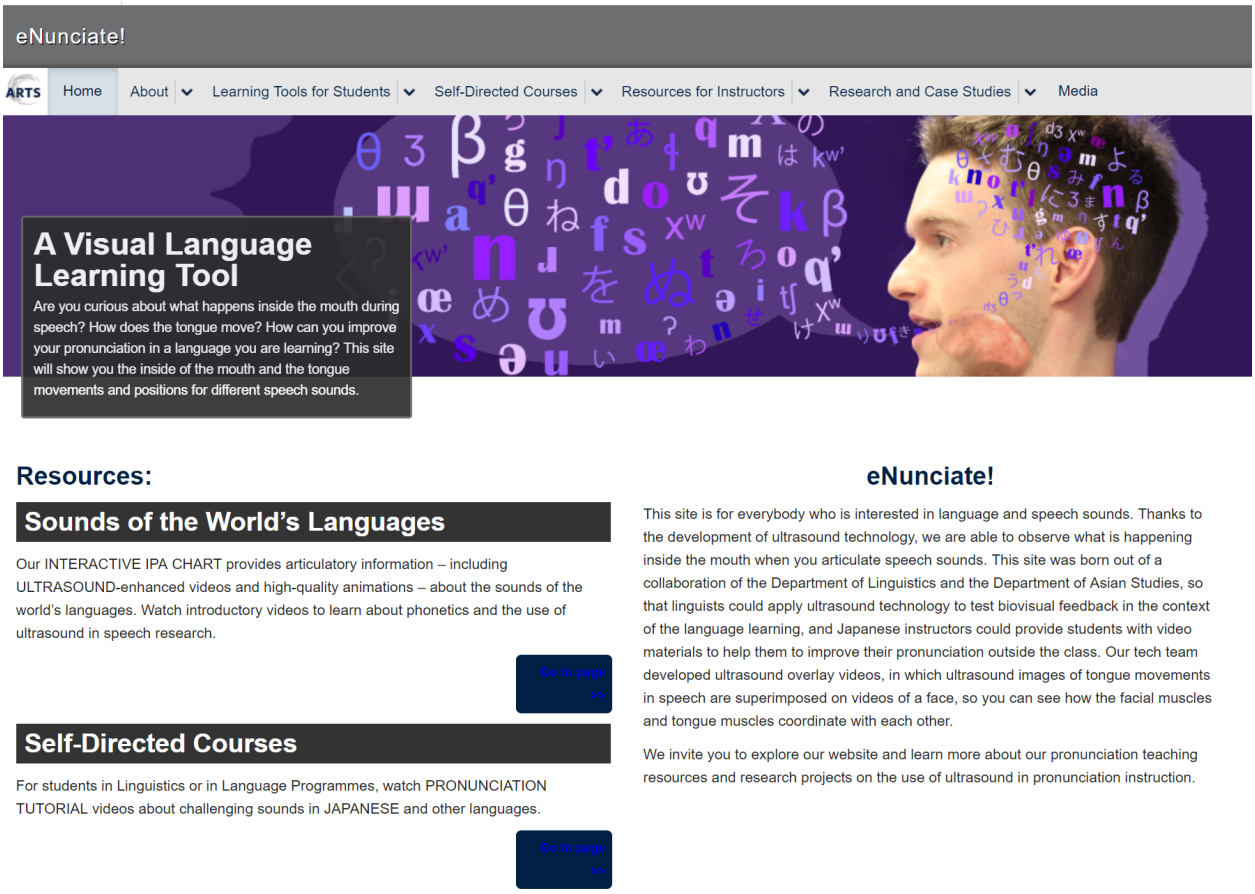


Figure 5. Screen shot of home page of eNunciate! Website.

For students, there are introductory videos as well as clickable IPA charts with real-time ultrasound imaging and with animations. See Figure 6 for sample consonant and vowel sounds.

Home / x

/x/ Voiceless Velar Fricative (Lower-case X)

Go Back to Chart

Instructions:

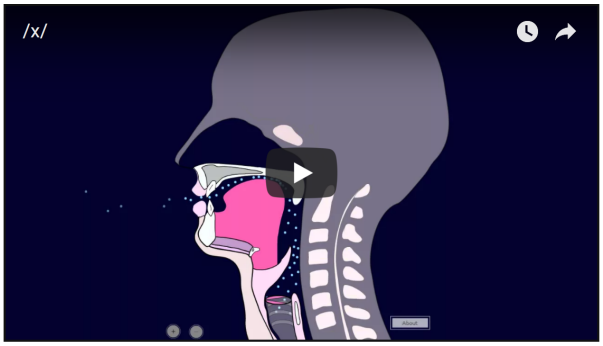
Articulator: Back of the tongue (tongue dorsum)

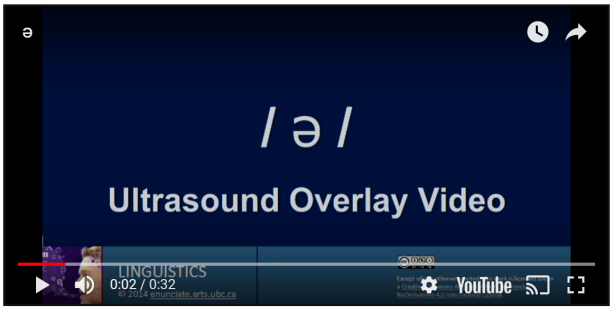
Point of Articulation: Velum

Manner of Articulation: Fricative: The articulator makes a narrow constriction against the point of articulation, so when the airflow goes through the oral passage, audible frication noise is created.

Voice: voiceless – Vocal folds are apart and are not vibrating.

Occurrence:
Abaza, Adyghe, Afrikaans (some speakers), Aleut (Atkan dialect), Arabic (Modern Standard), Assamese, Assyrian Neo-Aramaic, Avar, Azerbaijani, Basque (some speakers), Breton,





Source	Notes	URL
ArticulatoryIPA	Ultrasound	http://www.youtube.com/watch?v=kkOEKID3g7U
	Ultrasound – half-speed	http://www.youtube.com/watch?v=mEG3BN46eE0
		http://www.youtube.com/watch?v=wLXiiWEL8Qs
		http://www.youtube.com/watch?v=BN3M6HEZOaQ
Ken Stevens (1962)	X-Ray	http://www.youtube.com/watch?v=DcNMCB-Gsn8&list=PLDC5CFA68CD4AD861#t=0m11s
		http://linguistics.berkeley.edu/acip/course/chapter3/x_ray.html
University of South California	MRI; with noise	http://sail.usc.edu/span/videos/dai.mov
SpeechTrainer.de (Youtube)	Animation	http://www.youtube.com/watch?v=XvFdN1_3Bn4

Figure 6. Screen shot of real-time ultrasound imaging and animations of sample consonant and vowel.

There are Self-Directed Courses with tutorial videos about challenging sounds in Japanese and Cantonese with other languages still under development; additional tools like a Prosody Analyzer are likewise under development. For language instructors, there is also the Tongue Visualizer kit with which to produce custom resources that target language-specific pronunciation challenges. (eNunciate! Tongue Visualizer)

Seeing Speech (http://www.seeingspeech.ac.uk/), developed by phoneticians at the University of Glasgow, is another articulatory web resource for the study of phonetics, and includes MRI, ultrasound and animations (Figure 7). It provides clickable IPA charts (http://www.seeingspeech.ac.uk/ipachart/) with multiple visualizations by male and female talkers, and users can select by consonants or vowels; by MRI, ultrasound or animation; and by talker (not all sounds and tools available for both talkers).

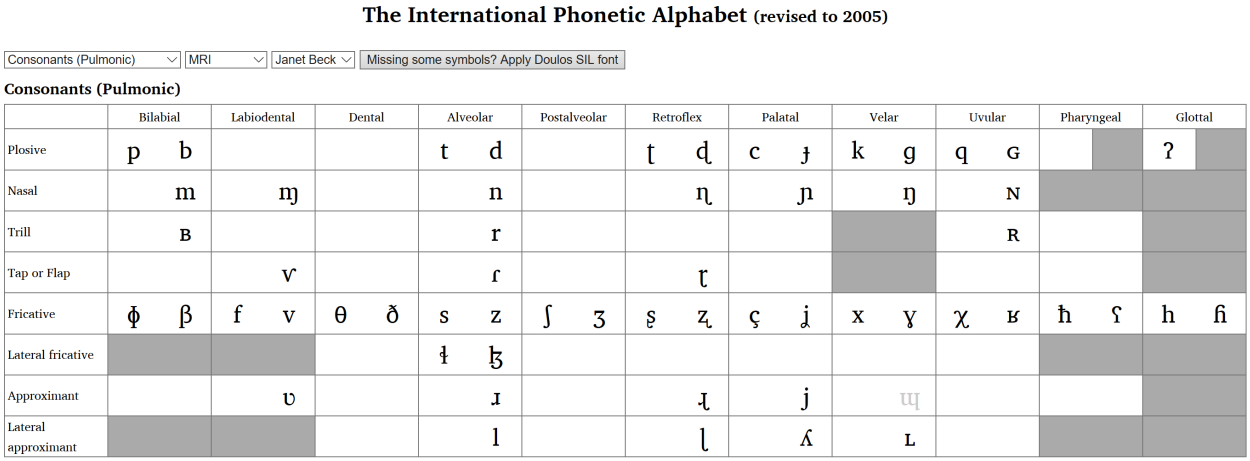


Figure 7. Screen shot of Seeing Speech's IPA clickable chart.

An additional resource linked to this site is **Dynamic Dialects** (http://www.dynamicdialects.ac.uk/), which contains a clickable world map and well as a clickable table (http://www.dynamicdialects.ac.uk/accentchart/accentchart.php) of speakers of different dialects producing a variety of words as well as extended spontaneous speech (last column of chart). Ultrasound tongue imaging of real-time captures is accompanied by inset videos of front and side views of the selected articulation.

Figure 8. Screen shot of female talker from Georgia pronouncing *lot* from Seeing Speech's Dynamic Dialects clickable chart.

CONCLUSION

Real-time visualizations of the articulation of speech are invaluable to improve understanding of the fine-grained mechanisms of speech production of one's native language, and they are very helpful in the learning of and practice with the pronunciation of non-native sounds and sequences. There are excellent online resources with MRI, ultrasound and animation imaging, and the availability and relative affordability of portable ultrasound devices further allows students and instructors to facilitate the development of the proprioception of different and challenging articulatory features as speakers develop their mastery of second language phonology.

ABOUT THE AUTHOR

D. Eric Holt is Associate Professor of Spanish and Linguistics in the Department of Languages, Literatures, and Cultures and the Linguistics Program at the University of South Carolina. His

scholarly interests include Hispanic linguistics, historical phonology, dialectology and phonological theory, as well as second language phonology, particularly of connected speech.

Department of Languages, Literatures, and Cultures, and Linguistics Program, University of South Carolina, 1620 College Street, Columbia, SC 29208, (803) 777-2063, deholt01@mailbox.sc.edu.

SELECTED ONLINE RESOURCES

http://enunciate.arts.ubc.ca/. eNunciate! University of British Columbia. (Ultrasound with Tongue Visualizer)

Research by the eNunciate Team listed at http://enunciate.arts.ubc.ca/research-and-case-studies/team-research/

Bird, S., H. Bliss, and B. Gick. (2017). Visualizing Speech in a Classroom Setting Using Interactive Ultrasound Imaging. 60-minute workshop presented at the 2017 PSLLT Research Methods Workshop. Salt Lake City, University of Utah: August 31, 2017. [slides] (this PSLLT conference pre-conference workshop)

http://sail.usc.edu/span/rtmri_ipa/index.html (MRI)

real-time MRI IPA charts. span | speech production and articulation knowledge group

http://www.seeingspeech.ac.uk/ipachart/ (MRI, ultrasound, animation)

Lawson, E., Stuart-Smith, J., Scobbie, J. M., Nakai, S., Beavan, D., Edmonds, F., Edmonds, I., Turk, A., Timmins, C., Beck, J., Esling, J., Leplatre, G., Cowen S., Barras, W., Durham, M. (2015). Seeing Speech: an articulatory web resource for the study of Phonetics. University of Glasgow. 1st April 2015. http://www.seeingspeech.ac.uk/

REFERENCES

Adler-Bock, M. (2004). *Visual feedback from ultrasound in remediation of persistent /r/ errors: Case studies of two adolescents*. MS thesis, University of British Columbia.

Adler-Bock, M., Bernhardt, B. M., Gick, B., & Bacsfalvi, P. (2007). The use of ultrasound in remediation of North American English/r/in 2 adolescents. *American Journal of Speech-Language Pathology*, *16*(2), 128-139.

Bernhardt, B., et al. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics & Phonetics 19*(6-7), 605-617.

Bernhardt, B., Bacsfalvi, P., Gick, B., Radanov, B., & Williams, R. (2005). Exploring electropalatography and ultrasound in speech habilitation. *Journal of Speech-Language Pathology and Audiology 29*, 169–182.

Bernhardt, B., Gick, B., Bacsfalvi, P., & Ashdown, J. (2003). Speech habituation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics and Phonetics 17*, 199–216.

Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In M. J. Munro & O.–S. Bohn (Eds.), *Second language speech learning: The*

role of language experience in speech perception and production. Amsterdam: John Benjamins.

Campbell, F. M. (2004). *The gestural organization of North American English /r/: A study of timing and magnitude*. MA Thesis, University of British Columbia.

de Bot, C. L. J. (1980). The role of feedback and feedforward in the teaching of pronunciation. *System, 8*, 35-45.

Gick, B., Allen, B. & Stavness, I. (2017). Speaking tongues are actively braced. *Journal of Speech, Language and Hearing Research*, 60(3), 494-506.

Gick, B., Bernhardt, B., Bacsfalvi, P., & Wilson, I. (2008). Ultrasound imaging applications in second language acquisition. In J. Hansen & M. Zampini (Eds.), Issues in Second Language Phonology. Cambridge, U.K.: Cambridge University Press.

Gick, B., Wilson, I., Koch, K., & Cook, C. (2004). Language-specific articulatory settings: Evidence from inter–utterance rest position. Phonetica 61, 220–233.

Kelsey, C. A., Woodhouse, R. J., & Minifie, F. D. (1969). Ultrasonic observations of coarticulation in the pharynx. *Journal of the Acoustical Society of America 46*, 1016–1018.

MacKay, I. R. A. (1977). Tenseness in vowels: An ultrasonic study. *Phonetica, 34*, 325–351.

Sonies, B. C., Shawker, T. H., Hall, T. E., Gerber, L. H., and Leighton, S. B. (1981). Ultrasonic visualization of tongue motion during speech. *Journal of the Acoustical Society of America 70*, 683–686.

Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning, 15*(5), 441-467.

Pillot-Loiseau, C., Kamiyama, T., & Antolík, T. K. (2015). French /y/-/u/ contrast in Japanese learners with/without ultrasound feedback: Vowels, non-words and words. In *International Congress of Phonetic Sciences (ICPhS)* (pp. 1-5). International Phonetic Association: London.

Wilson, I. (2003). An ultrasound study of articulatory tongue gestures in bilingual Japanese-English children's speech. Paper presented at the 4th International Symposium on Bilingualism (ISB4), Arizona State University, Tempe, AZ.

Wilson, I., & Gick, B. (2006). Ultrasound Technology and Second Language Acquisition Research. In Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006), ed. Mary Grantham O'Brien, Christine Shea, and John Archibald, (pp.148-152). Somerville, MA: Cascadilla Proceedings Project. http://www.lingref.com/cpp/gasla/8/paper1497.pdf