



2016-06-01

# The Effectiveness of Using Written Feedback to Improve Adult ESL Learners' Spontaneous Pronunciation of English Suprasegmentals

Chirstin Stephens  
*Brigham Young University*

Follow this and additional works at: <https://scholarsarchive.byu.edu/etd>

 Part of the [Linguistics Commons](#)

---

## BYU ScholarsArchive Citation

Stephens, Chirstin, "The Effectiveness of Using Written Feedback to Improve Adult ESL Learners' Spontaneous Pronunciation of English Suprasegmentals" (2016). *All Theses and Dissertations*. 6062.  
<https://scholarsarchive.byu.edu/etd/6062>

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu).

The Effectiveness of Using Written Feedback to Improve Adult  
ESL Learners' Spontaneous Pronunciation  
of English Suprasegmentals

Chirstin Stephens

A thesis submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Master of Arts

Lynn E. Henrichsen, Chair  
Norman W. Evans  
K. James Hartshorn

Department of Linguistics and English Language  
Brigham Young University

June 2016

Copyright © 2016 Chirstin Stephens

All Rights Reserved

## ABSTRACT

### The Effectiveness of Using Written Feedback to Improve Adult ESL Learners' Spontaneous Pronunciation of English Suprasegmentals

Chirstin Stephens

Department of Linguistics and English Language, BYU  
Master of Arts

This report describes a systematic procedure designed to give students pronunciation feedback on suprasegmental features of English in spontaneous production (rather than students' pronunciation during a read-aloud task). The procedure was developed to find out if written feedback (given frequently enough) could impact students' spontaneous production of suprasegmentals. Pronunciation feedback was given to the treatment group by marking transcripts of spontaneous speech with written symbols. Both the treatment group and the control group received form-focused pronunciation instruction. After 14 weeks, there was no significant difference between the groups, but there was a statistically significant improvement in students' comprehensibility overall (regardless of the feedback condition). Students were also surveyed to determine if either group perceived a greater benefit from the pronunciation instruction or if either group perceived a greater improvement in pronunciation. Surveys revealed a meaningful correlation between the group that received the treatment and the group that found the pronunciation instruction to be beneficial.

*Keywords:* pronunciation feedback procedure, suprasegmentals, written feedback

## ACKNOWLEDGEMENTS

We truly stand on the shoulders of giants, and I am deeply grateful for the giants who have gone before me and inspired this research. I am also indebted to the countless mentors and faculty at BYU who have helped me learn and accomplish what I have, especially including the faculty on my thesis committee. The administration, staff, teachers, and students at Brigham Young University's English Language Center have helped not only with this research, but also with my personal growth and education throughout my time in the MA program. My cohort has been a great support as well, and many of them have read countless pages of this manuscript to offer feedback and encouragement. Friends, roommates, and ward members who have shown an interest and offered encouragement will never know how much it helped. I would like to acknowledge that my family was instrumental in my being able to finish this research and manuscript. There are no words to express my gratitude for their encouragement, time, help, love, support, and patience. Finally, it would not be possible to even fathom starting a project like this without Heaven's tender mercies and the grace of God to strengthen me. My sincerest gratitude goes out to the seemingly endless numbers of people who have made this project as well as my education possible. This project symbolizes for me how high you can ascend and how much you can see as you stand on great shoulders and are supported by people who love you.

## TABLE OF CONTENTS

Title Page .....	i
Abstract .....	ii
Acknowledgements .....	iii
Table of Contents .....	iv
List of Tables .....	vi
List of Figures .....	vii
Preface .....	viii
Introduction .....	1
Literature Review .....	2
The Case for Instruction .....	2
The Case for Corrective Feedback .....	3
Instruction Combined with Feedback .....	6
Previous Feedback Models .....	7
Research Questions .....	10
Methodology .....	10
Participants .....	11
Instructional Design .....	11
Form-focused instruction .....	12
Student recordings .....	12
Measuring Improvement .....	12
Intervention .....	12
Listener ratings .....	14

Dynamic Written Corrective Feedback .....	15
Results.....	17
Listener Ratings .....	17
Student Surveys .....	19
Discussion.....	20
Implications.....	21
Limitations .....	22
Future Research .....	23
Conclusion .....	24
References.....	26
Appendix A.....	30
Appendix B.....	31

## LIST OF TABLES

Table 1 .....	11
Table 2 .....	18
Table 3 .....	19
Table 4 .....	19

LIST OF FIGURES

Figure 1 ..... 13



## PREFACE

In accordance with TESOL MA program guidelines, this thesis was prepared as a manuscript to be submitted to the *Journal of Second Language Pronunciation*. This journal was selected because of its focus and audience. The *Journal of Second Language Pronunciation* publishes research based on pronunciation specifically, whereas other journals in language teaching and learning typically have a broader focus. Readers of the *Journal of Second Language Pronunciation* may find this article useful while designing their own pronunciation research, and this article concludes with several recommendations for future research.

Manuscripts that are submitted to the target journal should (1) be prepared according to the Publication Manual of the American Psychological Association 6th Edition and (2) contain approximately 8,500 words. This manuscript was prepared in accordance with both of these requirements. The final draft of the manuscript has 7,901 words.

Alternative target journals include *Language Learning* and *System*. While neither journal is exclusively dedicated to pronunciation research (and the audience for these journals isn't entirely comprised of individuals interested in pronunciation pedagogy), this article draws from research on grammar instruction and corrective feedback, which are very current issues in language teaching. Length requirements vary between these two target journals (articles submitted to *Language Learning* must contain fewer than 10,000 words, whereas articles submitted to *System* must have fewer than 7,000 words).

## Introduction

Feedback "...is one of the most powerful influences on learning and achievement" (Hattie & Timperley, 2007, p. 81). Feedback may be especially powerful in the case of pronunciation improvement because many ESL students have difficulty noticing the particular aspects of English they are mispronouncing on their own (Derwing & Munro, 2005; Harmer, 2007). Feedback can help learners notice their errors, which can aid acquisition (Counselman, 2015).

Due to the impact feedback can have on learning in general, researchers have investigated different feedback techniques (Ellis, 2001; Hattie & Timperley, 2007). Pronunciation feedback techniques that have been investigated include those in which feedback is generated using technology (Anderson-Hsieh, 1992; Cucchiaroni, Neri, & Strik, 2009; Engwall & Bälter, 2007) as well as techniques in which the feedback is generated by teachers (Ellis & Sheen, 2006; Walker, 2009; Saito & Lyster, 2012; Dłaska & Krekeler, 2013). Technology-generated feedback can be provided to students frequently without demanding extra time from the teacher, but this feedback may be difficult for learners to understand and implement or it may be based on pre-recorded sentences (some of which may be markedly dissimilar from spontaneous production). Additionally, feedback should be based on what L2 learners actually need rather than what a computer can provide, and individualized feedback (that is easy to understand and implement) is difficult to generate using technology. In contrast to technology-generated feedback, teachers can give feedback that is easy to understand and relevant to student's needs; however, teacher feedback can be very time-consuming and consequently, teachers are not able to give it as frequently as technology-mediated feedback can be given to students.

This article reports on the development, implementation, and effectiveness of a procedure designed to provide clear, frequent pronunciation feedback on ESL learners' spontaneous oral production of English. The procedure developed in our research borrowed elements from two other procedures: one described by Celce-Murcia, Brinton, and Goodwin (2010) and one developed by Hartshorn, Evans, Merrill, Sudweeks, Strong-Krause, and Anderson (2010). Our feedback procedure was tested with 29 intermediate to advanced adult learners in an intensive English program to determine whether or not the feedback treatment led to gains in pronunciation improvement. The learners were also surveyed to determine if they felt the procedure was effective.

### **Literature Review**

To help readers better understand the creation of our pronunciation feedback procedure (which includes explicit instruction combined with feedback), this section will review previous research that pertains to the value of explicit instruction, the value of feedback, and various feedback models. Second language grammar research will also be included in this section due to a lack of empirically-tested pronunciation feedback models as well as the success that some grammar feedback procedures have experienced.

### **The Case for Instruction**

Form-focused instruction refers to instructional activities that are designed to focus learners' attention on linguistic forms (Ellis, 2001; Spada, 1997). There has been considerable debate about the value of form-focused instruction because teachers have noted that the forms they have taught are still executed inaccurately at least some of the time (Couper, 2006; Brown & Larson-Hall, 2012) which can lead to feelings of frustration. Notwithstanding, Brown and Larson-Hall (2012) encourage teachers not to abandon instruction because "[i]n the field of SLA,

research has shown that learners are worst at learning information they are not taught” (p. 162). Even in situations where students are still producing nontarget-like forms after instruction, instruction helps learners notice less salient elements of the target language (Lyster, 2004) that they might not otherwise notice.

Not only can explicit instruction benefit students generally, it has also been shown to improve pronunciation (Saito & Lyster, 2012; Dłaska & Krekeler, 2013; Derwing & Munro, 2005) and many researchers recommend that instruction should focus on suprasegmental (or prosodic) sounds and those segmental sounds that are important for communication (Derwing & Munro, 2005; Derwing et al., 1998; Miller, 2006; Celce-Murcia et al., 2010). Suprasegmentals are important to teach explicitly because they are not represented transparently in the written code (Derwing & Munro, 2005; Gilbert, 1994; Celce-Murcia et al., 2010) and learners are less likely to notice them on their own. Three suprasegmentals were investigated in our research: sentence stress, intonation, and rhythm. Sentence stress was chosen because it has been found to be highly indicative of target-like pronunciation proficiency (Ma, 2015; Kang, 2010). Rhythm and intonation were chosen because we felt they could be easily targeted in classroom instruction.

### **The Case for Corrective Feedback**

Accuracy and skill acquisition are concerns for students as well as teachers. Concerning writing, Ferris (1999) states that “[s]tudents themselves are very concerned about accuracy” (p. 1), and Harmer (2007) notes that most students expect feedback. Students specifically seek out feedback because they understand that feedback is necessary “...to reject or modify their hypotheses about how the target language is formed or functions” (p. 266). Regarding pronunciation, similar student expectations have been observed (Alghazo, 2015).

Even though experience has shown that students want feedback, whether or not to give feedback to L2 learners has been controversial. This controversy is highlighted by Ferris (1999), who, speaking of L2 writing, stated that “[r]esponding effectively to students’ grammatical and lexical problems is a challenging endeavor fraught with uncertainty about its long-term effectiveness” (p. 1). Truscott spoke out against error correction in both written (1996) and oral (1999) production, stating that error correction is not only ineffective, but potentially harmful. He argued that correction should be abandoned because it is difficult for teachers to give and doesn’t seem to impact students’ production. He also stated that students in writing classes should not spend valuable time on grammar at the expense of organization and rhetoric (1996). Truscott (1999) also wrote that feedback on oral production should be abandoned, stating that teachers may give feedback inconsistently (which is confusing for students), students may be negatively affected by the feedback, and students may not be ready for the feedback. Ferris (1999) and Lyster, Lightbrown, and Spada (1999), however, countered Truscott’s claims, offering encouragement to teachers to keep correcting student errors. Lyster, et al. (1999) agree that giving feedback is challenging, but state that it is feasible and even necessary in some instances. They further counter that feedback can be integrated into classroom activities skillfully to avoid Truscott’s concerns. The debate surrounding the topic of error correction will certainly continue to be debated until various points are settled by future research.

One convincing reason for giving feedback is the role it plays in proceduralization. In order to better understand the connection between feedback and pronunciation improvement, the role of proceduralization in skill acquisition theory will be briefly summarized in this section (for an extended explanation, see DeKeyser, 2014). According to Lyster (2004), “Skill development depends on transforming declarative representations, through practice, into production rules that

represent procedural knowledge” (p. 401). This transformation is referred to as proceduralization, and happens as students practice and receive feedback (Ellis, 2001; Lyster 2004). Both feedback and practice are “...crucial elements in information-processing models of L2 development because they engage learners in processes of restructuring interlanguage representation” (Lyster, 2004, p. 401). As learners restructure what they know about language, they can rely on feedback to inform their interlanguage rules.

Recent research that found a positive effect from feedback on pronunciation is reported by Dłaska and Krekeler (2013). They set out to determine if feedback was necessary for students to improve their pronunciation or if input alone was sufficient to impact pronunciation. Participants in the study recorded their speech (in a read-aloud task) and then participated in different treatments. One group participated only in listening activities after they recorded their speech in which they compared their recording to their teacher’s recording. The other group participated in the same listening activity (comparing their recording to their teacher’s), but they also received individual feedback. The comprehensibility of a pretest and a posttest sample for each participant (n=169) was rated. Those students who had received feedback were found to be more comprehensible at the end of the treatment.

While studies that integrated form-focused instruction with feedback have reported gains in students’ accuracy (both relating to grammar and pronunciation), not all have consistently reported improvement. According to Ellis (2001), those studies that have examined the effect of feedback on *written* accuracy have produced more conclusive results than studies that examine the effect of feedback on *spoken* accuracy. One example of research that didn’t find improvement in oral production (regardless of feedback condition) was cited by Lyster (2004).

While the research previously noted argues in favor of feedback because it addresses students' concerns and aids proceduralization, the value of feedback is nevertheless still debated from an empirical standpoint and merits further investigation, especially regarding the effectiveness of feedback on oral production.

### **Instruction Combined with Feedback**

Feedback cannot be divorced from instruction. According to Hattie and Timperley (2007), "Feedback can only build on something; it is of little use when there is no initial learning or surface information" (p. 104). They further assert "[f]eedback has no effect in a vacuum; to be powerful in its effect, there must be a learning context to which feedback is addressed" (Hattie & Timperley, 2007, p. 82). Larsen-Freeman (2014) also speaks in favor of combining instruction and grammar feedback stating, "[d]etailed instruction with explicit grammatical feedback may be the most helpful response to student errors" (p. 266).

While much of the research supporting instruction paired with feedback relates directly to grammar, in pronunciation research there have been similar findings. Saito and Lyster (2012) investigated corrective feedback in combination with form-focused instruction for pronunciation improvement. They examined the production of one segmental sound before and after instruction with three experimental groups (65 total participants). They found that the group who received meaning-based instruction with pronunciation instruction improved more than the group who received meaning-based instruction only. The group who improved the most received both types of instruction and pronunciation feedback, substantiating claims that form-focused instruction is most effective when accompanied by corrective feedback.

There remains, however, a great deal yet to learn about pronunciation instruction paired with feedback where both the instruction and the feedback focus on suprasegmental sounds.

## Previous Feedback Models

As the value of instruction and feedback is still debated and needs to be investigated further, we determined to include both explicit instruction and feedback in our study to determine the effect that feedback could have on suprasegmental production. When determining what our feedback procedure would be, we examined various techniques that have been used, which will be briefly described in this section.

In order to develop an effective feedback procedure, it is important to establish what effective feedback is. First, we should note that “...haphazardly correcting errors, either on oral language or written production, is not an effective way for a teacher to spend time with their students” (Brown & Larson-Hall, 2012, p. 107). Feedback should instead be deliberate, conscious, and sustained over a period of instruction. Hattie and Timperley (2007) describe effective feedback by stating that it should be “...clear, purposeful, meaningful, and compatible with students’ prior knowledge” (p. 104). In addition to some of the qualities already mentioned, Hartshorn et al. (2010) state that feedback should also be timely.

ESL teachers offer oral correction of pronunciation errors in class in various ways: by using recasts, elicitations, clarification requests, and explicit correction (Ortega, 2009). However, these types of in-class oral feedback may not lend themselves to being constant and sustained equally for each student. Another type of feedback that has been given is technology-mediated feedback. Anderson-Hsieh (1992) described a feedback procedure used with Chinese L1 teaching assistants. The procedure was developed to help the participants become more aware of typical English intonation patterns. In the procedure used by Anderson-Hsieh (1992), participants compared a visual representation of their intonation (produced by Visi-Pitch and the IBM Speech Viewer) against that of a native speaker’s pronunciation. Students first recorded



their own version of a sentence. Then, they looked at the pitch/intensity graph for their rendering next to the pitch/intensity graph of the same sentence pre-recorded by a native speaker. Students then re-recorded their utterance until the native speaker and nonnative speaker graphs were visually similar. Anderson-Hsieh's (1992) objective was to *describe* how electronic visual feedback was used in her teaching context, not to *evaluate* its effectiveness. While the effectiveness of her procedure was not evaluated, she did offer some observations about its perceived value and limitations. According to Anderson-Hsieh (1992), the feedback was beneficial because it helped students visualize intonation, was given in real time, and lowered student's self-consciousness. She noted, however, that students still would need to practice, monitor their own speech, and be able to transfer the feedback they received to communicative language production in order to achieve target-like production.

Celce-Murcia et al. (2010) recommend using audio-recordings as a way to give feedback on pronunciation. Based on their experience (but not any empirical research studies to our knowledge), Celce-Murcia et al. (2010) suggest that after teachers have set a clear time limit for open-ended speaking assignments, learners record their answers to a prompt, listen to what they recorded, and transcribe it. Then teachers listen to the recording, mark pronunciation errors on the student-produced transcript, and record their feedback (or a reformulation of the student's recording) in an audio file to send to students. Students then listen to the teacher's feedback/reformulation while looking at the marked-up script. Following all of these steps seems overwhelming for teachers to use frequently, as making teacher recordings in addition to marking transcripts can be time-consuming. For practitioners who use this procedure, Celce-Murcia et al. (2010) offer two cautions: (1) the task must have clear directions and (2) the teacher needs an organized, manageable way to give feedback.

Walker (2005) describes an adaptation of the above activity in which teacher recordings are not used to respond to pronunciation errors. In Walker's model, students read a monologue or a dialogue and then the teacher simply marks the script for them without recording feedback in an audio file. The effectiveness of Walker's model was evaluated by students (n=89). Students reported that they found Walker's model helpful. Nevertheless, students' actual progress was not measured. If proven effective in other studies, Walker's adaptation of removing teacher recordings may make oral feedback more manageable for teachers to give. Walker's model was not adopted for our research, however, because the students were not producing spontaneous speech, and our goal was to improve spontaneous production of suprasegmentals.

It is clear that each of these feedback techniques has different strengths. In-class feedback can be individualized to students, but may not be equally sustained for each student (or sustained across the semester). Electronic visual feedback could be sustained equally for each student, but relies on read-aloud tasks rather than spontaneous speech. Celce-Murcia et al. (2010) and Walker (2005) offer models in which the teacher could give feedback to students on spontaneous production, but the models (1) lack empirical research support and (2) don't provide for frequent feedback.

Pronunciation research that has studied frequent, teacher-generated feedback is altogether missing from pronunciation feedback studies. This is not surprising, owing to the increased demands that frequent, teacher-generated feedback adds to teachers' responsibilities; however, these studies are needed to determine if frequency is a major factor in the effectiveness of pronunciation feedback. Frequent grammar feedback has been studied, and has been shown to be effective (Hartshorn et al., 2010). Hartshorn et al.'s (2010) model (referred to as DWCF, or Dynamic Written Corrective Feedback) requires students to write a ten-minute paragraph during

each class session, which the teacher marks that day and returns to the students during the following class session. Students then revise their paragraphs, tally their errors, and resubmit their paragraph until all of the grammar errors have been resolved.

For our study, elements of Hartshorn et al.'s (2010) model were used in combination with some procedures recommended by Celce-Murcia et al. (2010) to design a pronunciation feedback procedure that could give frequent, meaningful feedback based on spontaneous suprasegmental production.

### **Research Questions**

This study investigated the impact that written feedback (without teacher recordings) can have on pronunciation production. If this procedure proved effective, it would be clearly advantageous, as the task of giving only written feedback would be less time-consuming than giving written feedback in addition to making teacher recordings. Our research investigated the answers to the following questions:

1. In the absence of teacher recordings, does written feedback combined with form-focused instruction lead to significant gains (as judged by human raters) in the spontaneous production of sentence stress, rhythm, and intonation for adult ESL students?
2. After receiving written feedback combined with form-focused instruction, do adult ESL students perceive a significant increase in their ability to spontaneously produce target-like sentence stress, rhythm, and intonation?

### **Methodology**

This section describes the participants in our study and the research design that we followed. This section also includes a brief comparison of the similarities and differences

between how DWCF is administered for improving written linguistic accuracy and which elements of DWCF were considered.

### Participants

Two listening/speaking classes comprising twenty-nine adult ESL students whose proficiency ranged from intermediate-high to advanced-low and were enrolled in an intensive English program participated in the study. Of these two classes, one was then randomly chosen to receive the treatment. The treatment group had 14 students and the contrast group had 15. These two classes were balanced by program administrators to make the members as similar as possible in L1 background and gender, as illustrated by Table 1.

Table 1

#### *Group Composition by Native Language and Gender*

Native Language	Experimental Groups					
	Treatment			Contrast		
	Male	Female	Total	Male	Female	Total
Spanish	2	5	7	4	4	8
Chinese	1	1	2	2	1	3
Korean	0	2	2	0	1	1
Japanese	0	1	1	1	0	1
Portuguese	0	1	1	1	1	2
Russian	1	0	1	0	0	0
Totals	4	10	14	8	7	15

Following standard disclosure procedures, and with the approval of the university's Institutional Review Board, participants received and signed informed consent forms.

### Instructional Design

Relative to this study, the two main elements of classroom instruction were (1) the form-focused pronunciation instruction that students received and (2) the recordings that students made. Each of these will be explained below.

**Form-focused instruction.** In an effort to limit the effect of teacher-related variables, the same teacher<sup>1</sup> (who was also the primary researcher) taught both groups using the same lesson materials and classroom practice activities, giving both groups (insofar as possible) the same amount of instructional time. For the majority of the semester, classroom instruction primarily focused on suprasegmentals, including sentence stress, intonation, and rhythm.

**Student recordings.** Both groups of students went to the computer lab twice a week to record their responses to a prompt (e.g., Describe your country). A new prompt was given for each session. Participants were given one minute to record their response. They were instructed to give their answers spontaneously (rather than writing it down and reading it), but they could practice their response before recording it. After each recording was created, students listened to their recording and transcribed it. Both the audio file and the transcript were submitted to the instructor.

### **Measuring Improvement**

Two different perspectives were considered when measuring the effectiveness of the intervention: the listener's and the student's. In order to contextualize both of these perspectives, this section will first describe the intervention. That description will be followed by a description of the procedures that were used to rate students' pronunciation improvement. The procedures that students used to rate their own improvement are also discussed.

**Intervention.** The goal of our study was to determine the role that written feedback plays in improving students' pronunciation. Nevertheless, rather than withholding feedback from the contrast group altogether, we determined to give the contrast group *grammar* feedback as a way

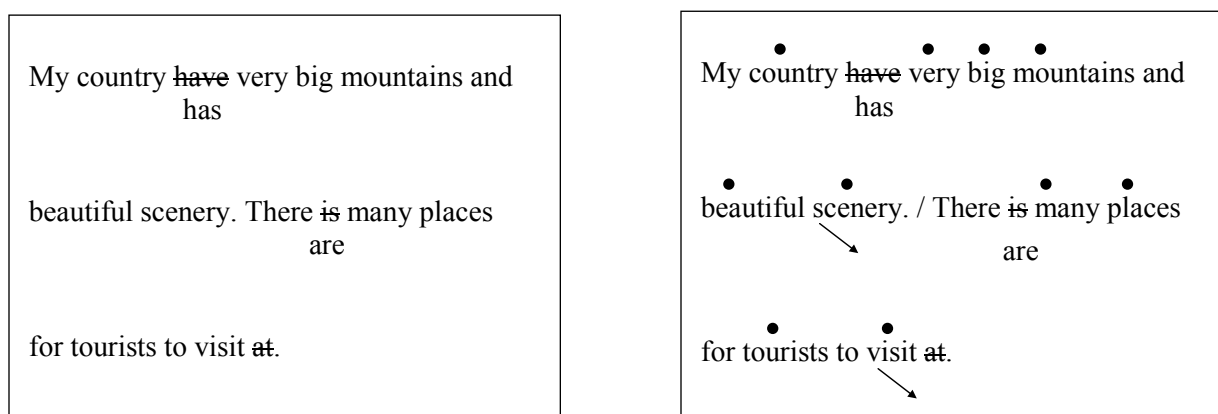
---

<sup>1</sup> The teacher was a TESOL MA student at the university connected with the English school where she taught. She had taken coursework in general TESOL methods as well as the teaching of listening, speaking, and pronunciation. She had also completed a TESOL internship and a practicum experience. Prior to this study, she had taught ESL for three years, including six pronunciation classes.

to respond to students' expectations of receiving feedback.. To keep all things as equal as possible between the two groups (except the focus variable), grammar feedback was also given to the *treatment* group. In addition to the grammar feedback, the instructor gave students in the treatment group written feedback about the accuracy of their suprasegmental pronunciation. Due to possible ambiguity and common usage in a variety of contexts, we will explain how written feedback was employed in our study. Students produced transcripts of their speech, which were marked by the teacher using symbols for different types of suprasegmental features of English. Dots were placed over words that should receive sentence stress, forward slashes were placed between words to indicate where students should pause, and arrows (either upward-pointing or downward-pointing) were drawn to indicate rising or falling intonation (respectively). These markings were explained in class and used in classroom instruction and practice activities (For an example of the written feedback used in this study, see Figure 1).

Figure 1

*Samples of feedback given to the contrast group and the treatment group*



The class session immediately following the creation of the audio recordings, the written feedback was given to the students and common errors were explained. Then students repeated the task, with the same prompt from the day before, 2-3 times in class, incorporating the

feedback they had been given by the instructor. They were encouraged to look at the feedback and ask questions about anything they didn't understand. They were then instructed to read the paragraph aloud one or two more times incorporating the feedback. The final repetition was from memory as much possible while students applied the feedback they had been given.

**Listener ratings.** Improvement was measured by rating and comparing one pretest and one posttest speaking sample for each participant. Pretest and posttest prompts were determined to be of approximately equal difficulty (See Appendix A for the full text of the prompts) and controlled for equal length (approximately 45 seconds). Each of the audio files was anonymized and assigned to two of the three raters.

These raters were TESOL MA graduates who were experienced teachers and raters. They were chosen because they had previous experience rating speech with an earlier version of the rubric that was used in this study. Still, the raters were re-calibrated to the rubric before completing the ratings for this study. Raters were unaware of the research questions for the study. Additionally, raters were not aware of which samples were pretest samples and which samples were posttest samples, nor were they aware of which samples were from students in the treatment group and which samples were from students in the contrast group.

Each rater was given a binder with a physical copy of the rubric for each sample they needed to rate. Raters were also given a USB drive with all of the audio samples they needed to rate. The files and paper rubrics were coded with a special number for each sample (that only the researchers were aware of) because numbering the rubrics and samples facilitated ordering them in such a way that rubrics and audio files were in the same sequential order.

The rubric used in this study was a modified version of Ma's (2015) pronunciation rubric. This rubric was selected for our study because it targeted the prosodic sounds we were

investigating and had been shown to be reliable (Ma, 2015). Two types of revisions were made to the rubric (in consultation with Ma). The first type was necessitated by the scope of the current study. Ma's original rubric included rating categories for vowels, consonants, word stress, sentence stress, intonation, and rhythm. Segmental sounds were excluded from the present study, so those categories were removed from the rubric. The second type of revision was primarily editorial and was intended to help clarify the rubric for the raters. The rubric used for this study is included in Appendix B.

**Student surveys.** In order to determine whether students felt they had improved in their pronunciation by the end of the course (our second research question), we had students complete a survey consisting of a series of questions about different aspects of the course. The first section asked about specific elements of pronunciation (i.e. sentence stress, intonation, and pausing). Students responded using a four-point Likert scale that ranged from 1 (*didn't improve very much*) to 4 (*improved a lot*). The second section asked students to circle which benefits they felt they had obtained from the course. To respond, from a list of 9 possible areas of improvement, they marked the benefits they felt they had gained from the course. They were also allowed the opportunity to write in any additional comments they had about the class.

### **Dynamic Written Corrective Feedback**

Where possible, principles from DWCF were implemented without adjustment into our procedure. For example, in both DWCF for written linguistic accuracy and for spoken pronunciation accuracy, students receive feedback the following class session that is individualized (feedback is based on students' spontaneous production). Students then need to attend to the individualized feedback by completing a reformulation of what they produced the day before as they incorporate the feedback they were given.



However, every aspect of DWCF was not applied to Celce-Murcia et al. (2010)'s procedure due to logistical constraints (e.g., time available to spend in the computer lab), manageability, and differences between grammar and prosody. The first adjustment we made to the DWCF model was the frequency with which students recorded paragraphs. Frequency was not mentioned by Celce-Murcia et al. (2010), but we realized that daily paragraphs as used in DWCF would not be manageable for pronunciation feedback. Our model required students to record oral paragraphs twice a week.

Another difference between DWCF and our model was the manner in which feedback was marked; namely, that in the framework established by Hartshorn et al. (2010), teachers identify and code errors with symbols without correcting the error for the students (e.g., the teacher would write "det" where a student made an error in either omitting a necessary determiner or using the wrong determiner). In our model, the teacher identified the errors explicitly by marking the word that should be prominent in a thought group, marking appropriate pauses, or drawing intonation arrows (rather than simply marking "int" to identify an intonation error). The decision to give direct feedback was made for several reasons. As previously noted, many ESL students have difficulty hearing their own errors in pronunciation, which means that the learners would need help understanding indirect feedback (at least some of the time). There was not enough time for the teacher to meet with each student about each paragraph, which would mean students would need to use other resources to understand their feedback. One such resource available at the school was a tutor lab. Using the tutor lab was not an ideal way for students to understand feedback because students were not likely to use it consistently and tutors were not trained in pronunciation pedagogy (which could have led to conflicting or otherwise confusing advice from the tutors).

The final major difference between DWCF and our model was the absence of a tally sheet or error log. In Hartshorn et al.'s (2010) model, students tally and keep track of each type of error they make. Errors weren't quantified in our model because there can be a great deal of individual variation in the prosodic features of native-like speech that were investigated in our study. This variation made it difficult to quantify. For example, some speakers tend to use shorter thought groups, which means that in their speech more words receive prominence (major stress) and there are more pauses. Some speakers tend to use longer thought groups, which means fewer words receive prominence and there are fewer pauses. This variation is problematic because there is no standard way to quantify how many pausing errors students exhibit. Sentence stress is another example of variability in prosodic features of spoken English. While there is no set number of words that should receive prominence in an utterance, at least some words should. The variability in acceptable stress placement makes quantifying errors difficult because many ESL students tend to speak with what some listeners call a "monotone" or "flat" intonation. Due to the variability of prosodic features in native-like speech, it was likely that if prosodic errors were to be quantified, they would not be quantified consistently, making an error tally sheet less effective than it was in Hartshorn et al.'s (2010) model.

## **Results**

### **Listener Ratings**

Facets ® software (Linacre, 2015) was used to analyze ratings and create fair averages for each of the samples based on the Many Facet Rasch Measurement (MFRM). Then we conducted a repeated measures Analysis of Variance (ANOVA). The measures for each of the pronunciation elements (e.g., comprehensibility, sentence stress, intonation, and rhythm) were

calculated separately with one between-subjects factor (group) and one within-subjects factor (time).

**Comprehensibility.** The pre-test and post-test comprehensibility score means (and standard deviations) for both groups are shown in Table 2. Values are based on a 7-point comprehensibility scale. The differences shown in the table did not seem to be very large, but the ANOVA revealed a main effect of Time in the comprehensibility ratings  $F(1,27) = 11.676$ ,  $p=.002$ ,  $\eta_p^2 = .302$ . This result indicated that there was significant improvement in comprehensibility for the participants generally from the pretest to the posttest. Nevertheless, there was no statistical significance between groups,  $F(1,27) = .196$ ,  $p=.661$ ,  $\eta_p^2 = .007$ , indicating that the treatment group did not improve any more than the contrast group.

Table 2

*Comprehensibility Pretest and Posttest Means and Standard Deviations by Group and Overall*

	N	Pretest		Posttest	
		M	SD	M	SD
Control Group	15	4.90	1.143	5.47	1.058
Treatment Group	14	4.78	1.067	5.22	1.087
Overall	29	4.84	1.089	5.35	1.061

**Suprasegmentals.** Separate ANOVA measures were collected for each of the suprasegmental categories. The categories of sentence stress, intonation, and rhythm did not show any significant differences for either Time or Group (See Table 3).

Table 3

*ANOVA Results*

	<i>F</i> (1,27)	<i>p</i>	$\eta_p^2$
Comprehensibility			
Time	11.676	.002	.302
Time x Group	.196	.661	.007
Sentence Stress			
Time	1.554	.223	.054
Time x Group	.465	.501	.017
Intonation			
Time	.927	.344	.033
Time x Group	.425	.520	.016
Rhythm			
Time	.398	.534	.015
Time x Group	.055	.817	.002

**Student Surveys**

Responses regarding the degree of perceived improvement in each area were analyzed using a univariate ANOVA, which found that there was no statistically significant main effect for the group by Area,  $F(8,234) = .162$ ,  $p=.969$ ,  $\eta_p^2 = .010$ . Responses regarding the perceived benefits of the course (tallied by frequency and shown in Table 4) revealed a meaningful correlation between the group that received the treatment and the group that found the pronunciation instruction to be beneficial ( $X^2 = 3.548$ ,  $p=.06$ ,  $\Phi=.35$ ).

Table 4

*Benefits of the Course Reported by Students*

	Pronunciation Instruction	Improved Pronunciation	Improved Grammar
Contrast	8	11	9
Treatment	12	8	9

A qualitative analysis of the open-ended responses in this section of the survey revealed that the students were pleased with the course and the teacher. However, none of the comments focused on the feedback procedure provided.

### **Discussion**

The goal of our research was to determine whether written feedback alone could improve students' spontaneous production of suprasegmentals. Insofar as the results obtained in this study can be generalized, written feedback alone (even if it is timely and constant) is insufficient to improve the spontaneous production of suprasegmentals for adult ESL learners. This finding may lend support to the aforementioned feedback procedures outlined by Celce-Murcia et al. (2010). However, there may have been other important moderating variables (motivation, time of day, length of the study, etc.) that also influenced the results noted.

While our *intervention* did not significantly impact students' pronunciation, *instruction* may have. The improvement noted for both groups from the pretest to the posttest may have resulted from the form-focused pronunciation instruction given as part of our study. This kind of improvement is somewhat rare in pronunciation studies. Rossiter, Derwing, Manimtim, and Thomson (2010) state that "learner improvement is unlikely to be [noted] over the duration of a single ESL course" (p. 600). The improvement shown by both groups should offer encouragement to practitioners who are unsure about whether or not pronunciation can improve in a relatively short amount of time. It should be noted that the improvement may also be attributed to other factors, and the cause of the improvement is difficult to determine. Students may have improved due to individual differences, varying personal situations, or maturation. According to Morley (1994), not very many students are able to acquire pronunciation satisfactorily relying on input, but it is possible that our participants' improvement came from

outside sources that could not be controlled for (e.g., native-speaking roommates, better hearing ability, different motivation, etc.).

Our study of students' perceptions revealed that more students in the treatment group (12 of 14) felt that the pronunciation instruction was valuable when compared to students in the contrast group (8 of 15). Students in the treatment group may have valued instruction more than students in the contrast group did because the instruction was directly linked to the feedback they received and was necessary to make the feedback more understandable. This possible interpretation is supported by Lyster (2004) and Hattie and Timperley (2007) that feedback must be related to previous instruction.

Students in the contrast group (11 of 15) indicated more frequently (compared to 8 of 14 in the treatment group) that they felt the class was beneficial because they improved their pronunciation. Perhaps because they did not receive teacher feedback, students in the contrast group were less aware of their suprasegmental error patterns than students in the treatment group and thus evaluated themselves more generously. However, differences in learning style, educational background, or perceptions about what "good" pronunciation instruction should include may also have influenced students' perceptions.

### **Implications**

In general, students will likely continue to expect help improving their pronunciation, and practitioners that give feedback under conditions similar to those used in our study should consider combining visual and auditory feedback, as written feedback alone may not provide students with enough linguistic information to implement the teacher's corrections. Of course, the effectiveness of this combination of feedback modes (as well as auditory feedback alone) would need to be investigated empirically.

Another important implication of our study is that increasing the frequency of feedback may not be as important as ensuring that students understand and know how to implement the feedback they receive. Teachers who give pronunciation feedback should take measures to ensure that students understand how to apply the feedback they are given.

As previously noted, some practitioners feel that explicit pronunciation instruction may do little to actually improve students' pronunciation even though accuracy is highly valued by students themselves. Notwithstanding the fact that our study did not find significant gains in students' spontaneous production of suprasegmentals, it did reveal that many students were able to make some gains (even marginal ones) in their comprehensibility over the course of a single semester, which is encouraging.

### **Limitations**

Perhaps the most fundamental limitation of using written feedback as executed in our study was the difficulty some students seemed to experience in knowing how to apply the feedback they were given. For example, even though students cognitively understood they needed to use falling intonation in a certain utterance, many were unable to incorporate the feedback accurately.

The manageability of our feedback procedure was another limitation in our design. As with most ESL classes that are not dedicated exclusively to pronunciation improvement, the time the teacher (who was also the primary researcher) spent giving pronunciation feedback was in addition to regular grading and planning for instruction. The constraints on the teacher's time limited the amount of feedback that could be given. If the teacher could have given the feedback more frequently, students may have been able to improve more. It is unclear, however, that more frequent feedback would have led to greater gains in pronunciation accuracy. As noted in the

first limitation, many students seemed to struggle to apply written markings to their pronunciation, and the quantity of feedback may not have improved its usefulness.

### **Future Research**

First, future research is needed to determine if written feedback could be delivered differently to make it more effective. Replications of this study that implement more frequent feedback or measures to make the feedback more understandable may help give a clearer picture of the potential written feedback has to improve pronunciation. To allow for an increase in the frequency of the feedback, future studies could be done in dedicated pronunciation courses where there are fewer demands on the teacher's time.

Another suggestion for future researchers is to investigate further ways to make feedback manageable. Steps can be taken to ensure that students understand the feedback they are given, but improving manageability remains problematic. Integrating standardized recorded readings with spontaneous recordings may be a way to address this concern. Students could get frequent feedback that is less time-consuming for teachers because teachers could record the same passage that students read once (rather than making a unique feedback recording for each student) and then make it electronically available to all of the students through a learning management system, teacher blog, email, etc.

More research is also needed to determine if auditory feedback (in accordance with the recommendation made by Celce-Murcia et al. (2010) noted previously) would make written feedback more accessible for students. A study in which written pronunciation feedback (similar to what we used in our study) is compared to written and audio feedback would be especially helpful in determining the effect written feedback can have on production.



To increase their chances for success, future studies should use a larger sample than we did and randomly assign students to groups rather than using intact classes. Studies in different teaching contexts would also offer additional insights. Teachers and/or researchers in other settings (especially EFL contexts) may find very different results with the same procedure.

### **Conclusion**

Many teachers and students are aware that feedback is a valuable tool to help students increase their accuracy. It may be especially important to find effective ways of helping students notice their L2 pronunciation errors, as many students are unable to hear them. Recommended techniques for L2 pronunciation feedback have a tendency to be very time-consuming for teachers, which tends to decrease the frequency of the feedback students receive. Dynamic Written Corrective Feedback (DWCF) for written linguistic accuracy as developed by Hartshorn et al. (2010) allows for frequent feedback. Our research integrated DWCF principles with pronunciation improvement techniques recommended by Celce-Murcia et al. (2010) in order to determine if more frequent feedback could have a greater effect on pronunciation without overwhelming teachers. Teacher audio recordings are recommended for giving feedback, but as they can be very time-consuming, they were eliminated in our model to make the feedback more manageable for teachers to give (and consequently, increase the frequency of the feedback that was given). Our investigation determined that while there was overall improvement in comprehensibility for both groups of learners after 14 weeks of form-focused instruction, there was not a significant difference in the production of suprasegmental sounds between the groups. Future research should determine the effectiveness of using teacher recordings combined with written feedback and, if proven effective, how to make that feedback more manageable for

teachers. Knowing that feedback is a powerful influence on learning is insufficient; we must also learn how to harness its power effectively.

## References

- Alghazo, S. M. (2015). Advanced EFL learners' beliefs about pronunciation teaching. *International Education Studies*, 8(11), 63-76. DOI: 10.5539/ies.v8n11p63
- Anderson-Hsieh, J. (1992). Using electronic visual feedback to teach suprasegmentals. *System*, 20, 51-62.
- Brown, S. & Larson-Hall, J. (2012). *Second language acquisition myths: Applying second language research to classroom teaching*. Ann Arbor, MI: University of Michigan Press.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. B. (2010). *Teaching pronunciation: A course book and reference guide* (2nd Ed.). New York, NY: Cambridge University Press.
- Counselman, D. (2015). Directing attention to pronunciation in the second language classroom. *Hispania*, 98(1), 31-46. DOI: 10.1353/hpn.2015.0006
- Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect*, 21(1), 46-66.
- Cucchiarini, C., Neri, A., & Strik, H. (2009). Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback. *Speech Communication*, 51(10), 853-863.
- DeKeyser, R. (2014). Skill acquisition theory. *Theories in second language acquisition: An introduction (2nd edition)*. In VanPatten B., Williams J. (Eds.), Florence, KY, USA: Routledge.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379-397.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48, 393-410.

- Dlaska, A., & Krekeler, C. (2013). The short-term effects of individual corrective feedback on L2 pronunciation. *System*, 41(1), 25-37. DOI: 10.1016/j.system.2013.01.005
- Engwall, O., & Bälter, O. (2007). Pronunciation feedback from real and virtual language teachers. *Computer Assisted Language Learning*, 20, 235-262.
- Ellis, R. (2001). Investigating form-focused instruction. *Language Learning*, 51 (Suppl. 1), 1–46.
- Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28, 575-600.
- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott. *Journal of Second Language Writing*, 8(1), 1-11.
- Gilbert, J. B. (1994). Intonation: A navigation guide for the listener (and gadgets to help teach it). In J. Morley (Ed.) *Pronunciation pedagogy and theory: New views, new directions* (pp. 36-48). Bloomington, IL: Pantagraph Printing.
- Goodwin, J. (2014). Teaching pronunciation. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Ed.) (4th ed.). *Teaching English as a second or foreign language* (pp. 136-152). Boston, MA: National Geographic Learning.
- Harmer, J. (2007). *The practice of English language teaching*. Harlow, England: Pearson Education Limited.
- Hartshorn, K. J., Evans, N. W., Merrill, P. F., Sudweeks, R. R., Strong-Krause, D., & Anderson, N. J. (2010). Effects of dynamic corrective feedback on ESL writing accuracy. *TESOL Quarterly*, 44, 84-109.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.

- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 301-315. DOI: 10.1016/j.system.2010.01.005
- Larsen-Freeman, D. (2014). Teaching Grammar. In M. Celce-Murcia, D. M. Brinton, & M. A. Snow (Ed.) (4th ed.). *Teaching English as a second or foreign language* (pp. 256-270). Boston, MA: National Geographic Learning.
- Linacre, J. M. (2015) Facets computer program for many-facet Rasch measurement, version 3.71.4. Beaverton, Oregon: Winsteps.com
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26(3), 399-432. DOI 10.1017/S0272263104263021
- Lyster, R., Lightbrown, P. M., & Spada, N. (1999). A response to Truscott's 'What's wrong with oral grammar correction'. *Canadian Modern Language Review*, 55(4), 457-467.
- Ma, R. (2015). *The Role of Pronunciation in Speaking Test Ratings*. All Theses and Dissertations. Paper 4426.
- Miller, S. F. (2006). *Targeting pronunciation: Communicating clearly in English* (2nd Ed.). Boston, MA: Houghton Mifflin Company.
- Morley, J. (1994). A multidimensional curriculum design for speech-pronunciation instruction. In J. Morley (Ed.) *Pronunciation pedagogy and theory: New views, new directions*. (pp. 64-91). Bloomington, IL: Pantagraph Printing.
- Ortega, L. (2009). *Understanding Second Language Acquisition*. New York, NY: Routledge.
- Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *Canadian Modern Language Review*, 66, 583-606.

- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/. *Language Learning*, 62, 595-633.
- Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, 30 (1). 73-87.  
doi:10.1017/S0261444800012799
- Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Truscott, J. (1999). What's wrong with oral grammar correction. *Canadian Modern Language Review*, 55, 437-456.
- Walker, R. (2005). Using student-produced recordings with monolingual groups to provide effective, individualized pronunciation practice. *TESOL Quarterly*, 39, 550-558.  
<http://doi.org/10.2307/3588495>
- Yule, G. & Macdonald, D. (1995). The different effects of pronunciation teaching. *IRAL: International Review Of Applied Linguistics In Language Teaching*, 33, 345-350.

## Appendix A

### Actual Pretest and Posttest Prompts

---

Pretest Prompt	Describe the events surrounding a presidential election or other government event in your country. What are the reasons for these events? How do people participate? What are things that a person would see or do if they were in your country during this event?
Posttest Prompt	Describe a holiday in your country that other countries do not celebrate. What is the reason for the holiday? How do people celebrate? What are things that a person would see, do or eat if they visited your country during that holiday?

---

## Appendix B

Ma's 2015 rubric as adapted for this study

	<b>Sentence stress</b>	<b>Rhythm</b>	<b>Intonation</b>
5	Sentence stress is almost always placed appropriately based on the speaker's communicative intent.	Stress-timed rhythm is used naturally and consistently.	A variety of intonation patterns effectively reflect the speakers' intent (e.g., questioning, apology, sarcasm, etc.)
4	Sentence stress is placed correctly most of the time, but is sometimes misplaced.	Stress-timed rhythm is employed naturally most of the time.	Intonation is employed effectively to express emotion, but a particular pattern may be overused.
3	Sentence stress is employed, but not always correctly (e.g., function words receive stress inappropriately).	Stress-timed rhythm is employed with effort but sometimes appears unnatural.	Intonation is usually correct but occasionally misleads listeners.
2	Sentence stress is uncommon or is often misplaced.	Rhythm is mostly syllable-timed, but occasionally demonstrates stress-timing.	Rising and falling intonation patterns are sometimes used appropriately but may impede understanding.
1	Sentence stress is hardly ever used to indicate key words in thought groups.	Rhythm is predominantly and strongly syllable-timed (i.e., very "choppy").	Intonation is used inappropriately and interferes with communication or is distracting.