THE EFFECTS OF SELF-MONITORING STRATEGY USE ON THE PRONUNCIATION OF LEARNERS OF ENGLISH

BY

SUE ANN INGELS

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Educational Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor Emeritus Gary Cziko, Chair
Professor Wayne Dickerson, Research Director
Assistant Professor Melissa Bowles
Adjunct Assistant Professor Laura Hahn

Abstract

Language learning strategy (LLS) research has provided a large body of evidence for the effectiveness of strategy-based instruction (SBI), though the evidence is very limited for pronunciation strategy instruction. For both general and pronunciation LLSs, most research has focused on identifying the strategies used by successful learners. When strategy instruction has been investigated, in most cases the strategies that were taught were not directly linked to specific tasks, learners were not observed using the strategies, and measures of strategy effectiveness often were holistic and did not reveal improvements in specific pronunciation features.

The goal of this study was to extend our understanding of the role of strategy use in L2 (second language) pronunciation learning by investigating the effectiveness of training future international teaching assistants (ITAs) to critically listen to, transcribe, mark corrections (annotate), and orally rehearse English suprasegmental features in their own speech. The suprasegmental features investigated were message unit boundaries, primary phrase stress, intonation, vowel reduction in content and function words, linking, word stress, and multiword construction stress. Fifteen graduate-level learners of English (14 Mandarin speakers, 1 Korean speaker) from an intact English as a Second Language (ESL) pronunciation class at a Midwestern university were solicited to participate in a repeated-measures design, in which the independent variables were 3 levels of self-monitoring (listening only [L], listening + transcription [LT], and listening + transcription + annotation [LTA]) and rehearsal (R). The strategies were examined in the following combinations: LR-LR-LR, LT-RRR, and LTA-RRR. The dependent variable was the change in suprasegmental accuracy following self-monitoring and rehearsal. Speech data resulting from strategy use were gathered at the beginning and end of

a 16-week semester in order to determine the extent to which strategy use corresponded to improved suprasegmental accuracy.

Key findings include the following: (a)All participants made meaningful improvements in suprasegmental accuracy for at least some of the targets following self-monitoring; (b) the LT-RRR combination was most effective for lower proficiency learners and LTA-RRR was most effective for higher proficiency learners; (c) starting proficiency and size of accuracy gains following self-monitoring were negatively correlated; (d) self-monitoring had differential effects on accuracy for the suprasegmental features, with message units, linking, and function words showing the greatest improvement;(e) and observation of individual task performance provided useful insights into how effectively adult L2 learners utilize self-monitoring strategies. Implications for language teaching and learning, limitations of the study, and future research opportunities are explored.

In memoriam
Frances Mary McCaughey Ingels

Acknowledgments

The kind and unwavering support of many colleagues, fellow graduate students, friends, and family members has sustained me throughout the process of completing my research and writing this dissertation. Here is where I officially thank all of you, as well as assure you that I will remain ever thankful for your encouragement and help.

First, I extend my extreme gratitude to my research director, Wayne Dickerson, who has been my phonology, teaching, and research mentor and good friend for the last 9 years; to my committee chair, Gary Cziko, who offered statistical support and invaluable feedback on my research design and data analysis; and to committee members Melissa Bowles and Laura Hahn, who have offered encouragement and brilliant insights that follow from their rich research and teaching backgrounds. You each provided a wonderful balance of firm and clear guidance, encouragement to pursue my research agenda, and a sharing of your research, teaching, and publishing expertise. I am grateful for your artful guidance.

Numerous teaching colleagues assisted in a very tangible and important way, by completing some of the early stages of speech transcription, which in turn expedited my finer grain transcription of each participant's speech samples: Nagham Awadallah, Tamara Brayak, Cheongsil Byun, Shawn Fitzpatrick, Dustin Kelly, Marshall Klassen, Shannon McCrocklin, and Rana Sawwan. Thank you for handling this important task while juggling your many coursework and teaching duties.

Statistical advice was provided by Dr. Younshin Chi, Professor Gary Cziko, Professor Fred Davidson, and Bill Messner, with ATLAS Stats/GIS/Social Science Data Support. Your guidance allowed me to proceed with confidence regarding the use and interpretation of the inferential statistics presented in this dissertation.

Thank you also to MATESL colleague, friend, and teaching and research collaborator Amanda Huensch, for assistance with data analysis; Rick Partin, for assistance collecting and carefully guarding participant consent forms; and pronunciation instructor extraordinaire Patricia Watts, for providing creative pedagogical insights regarding pronunciation strategy use.

The many international graduate students enrolled in the ten-plus sections of pronunciation courses that I taught over the last several years provided the inspiration for investigating this research topic. They worked diligently to gain the upper hand over the English phonological system, while challenged by the demands of course- and lab work, and while engaged in their own research—all in a second language and thousands of miles from their home countries. They were and still are very inspiring indeed.

Staff, classmates, and professors, and the familiar and friendly faces in the halls and elevators of the Foreign Languages and Education Buildings have each played their part in providing encouragement—you all "know the drill" and know the importance of encouraging those of us faithfully completing the many steps that bring us closer to doctoral degree completion. Thank you, one and all.

Finally, I wish to acknowledge the love and rock-solid support of my son, Jamie Mauck, and life partner, Dennis Cluver.

Contents

Chapter 1

Introduction

This study investigated the effectiveness of training adult L2 (second language) learners to critically listen to, transcribe, identify corrections for, and orally rehearse L2 suprasegmental features in their own speech. Collectively I refer to this set of strategies as "self-monitoring." Fifteen learners enrolled in a university ESL (English as a Second Language) pronunciation class participated in a repeated-measures two-factor design, in which three levels of self-monitoring and three levels of rehearsal were the independent variables. The dependent variable was the proportion of nontarget pronunciation features in an original speech sample that was corrected following self-monitoring and rehearsal.

Learners self-monitored eight target suprasegmental features in their oral production: message unit boundaries, primary phrase stress (also called prominence, focus, emphasis), intonation, vowel reduction in content and function words, linking, word stress, and multiword construction stress. The teaching context was an intact graduate-level ESL pronunciation class, for which a primary instructional goal was training learners to improve their English pronunciation and develop self-study skills that could be used independently to further increase pronunciation accuracy, without requiring access to specialized technology or resources not commonly available to most learners.

Motivations for the Study

Learning to perceive and produce L2 sounds and prosody, and understanding how they function and are implemented in the L2, are important skills for adult language learners who want to maximize their spoken L2 intelligibility and continue to improve their oral skills outside

of classroom instruction. Improving language skills is an ongoing process and having access to tools that allow language learners to take control of their own L2 pronunciation learning are considered essential for achieving one's goals for academic and professional success in contexts where the L2 is required (Celce-Murcia, Brinton, & Goodwin, 2010; W. B. Dickerson, 1994; Morley, 1991; and others).

A variety of factors influence how successful adults are in learning an L2 phonology. Some learners acquire sufficient pronunciation accuracy without explicit instruction (Riney & Flege, 1998). Learners in this group typically learn an L2 before puberty or are gifted adult language learners. A number of factors influence adult L2 pronunciation acquisition, including factors outside the learners' control, such as language aptitude, phonemic coding ability, developmental readiness, and working memory (Celce-Murcia et al., 2010; Juffs & Rodriguez, 2007), as well as factors that learners have some ability to control, such as motivation and amount of L2 exposure, instruction, and use (Lightbown & Spada, 2006). However, for the learners within the context of the current study, motivation, daily exposure to English, or prior classroom instruction have not been sufficient for acquiring the necessary accuracy. These learners often are unable to identify when their pronunciation is not targetlike and do not notice relevant L2 features in native speaker (NS) speech, though they know their speech differs from target L2 production in ways that negatively affect intelligibility.

For more than 30 years, language teachers and researchers have been particularly interested in identifying what sets apart "good" language learners (those who are successful L2 speakers) from those who seem unable to make a desired level of L2 progress in spite of instruction, exposure, and their own language learning efforts (Oxford, 1990; Rubin, 1975, 1981; Stern, 1975; Vann & Abraham, 1990). Relevant questions have included the following: What are

the characteristics and language learning behaviors of successful learners? Can (and should) these behaviors (i.e., language learner strategies [LLSs]) be taught to all language learners to help ensure the most efficient and successful L2 acquisition? Does L2 proficiency level correspond in any way to the types of LLSs used and learners' success in using LLSs? To identify what language learners actually do to further L2 learning, LLS researchers have observed, surveyed, interviewed, and administered verbal self-report protocols to language learners. The outcome of this intensive and long-term study has been the identification and categorization of the strategies typically used by learners for the four primary language skills: speaking (i.e., general oral skills), listening, writing, and reading, and less frequently, test-taking skills (McDonough, 1995). LLS research results suggest a positive relationship between LLS use (quantity, frequency, and quality of strategy use) and proficiency (Breun, 2001b; Chamot & Kupper, 1989; Chesterfield & Chesterfield, 1985; Lai, 2009; Oxford & Nyikos, 1989). The question of the effectiveness of strategy-based instruction (SBI) still needs further investigation, though evidence from SBI research of the past two to three decades indicate that SBI typically is effective (Chamot & Kupper, 1989; Y. M. Chen 2008; Cohen, Weaver, & Li, 1995; Hassan et al. 2005).

Though a substantial body of research and scholarly discussion exists regarding strategies for improving general L2 skills and to a lesser extent, L2 speaking, until recently little attention has been focused on specific strategies and techniques that learners can use to improve L2 pronunciation (Eckstein, 2007; Haslam, 2010; Sardegna, 2009). In my own teaching and research, I have observed the value of training adult L2 learners to use self-monitoring strategies such as critical listening and transcription to monitor their output and self-correct their nontarget pronunciation. Early on, I wondered how accurately learners could evaluate their own

pronunciation and whether the effort required to listen to and transcribe one's own speech was worthwhile. Writers from at least 30 years ago have agreed on the need for self-monitoring for improving general L2 speaking abilities:

> The good language learner monitors his own and the speech of others. That is, he is constantly attending to how well his speech is being received and whether his performance meets the standards he has learned. Part of his monitoring is a function of his active participation in the learning process. He is always processing information whether or not he is called on to perform. He can learn from his own mistakes. (Rubin, 1975, p. 47)

W. B. Dickerson (1989, 2000) emphasizes the importance of helping L2 learners develop "self-critical abilities" (p. xiii) and implement systematic "covert rehearsal" (p. xvii) specifically for pronunciation improvement. However, in the past two decades limited additional work has been completed regarding how language learners implement self-monitoring strategies for improving pronunciation and indeed whether such strategies enable learners to make their pronunciation more targetlike. The current study was designed to address these gaps.

Contents of This Dissertation

This document consists of the following chapters: Chapter 2, Literature Review, which offers a discussion of how the pronunciation learning strategies of listening, transcription, correcting a transcript, and rehearsing corrections aloud are related to the larger literature on LLSs. Also discussed are the rationale for using these strategies and how learners can be taught to use them, the status of research on the effectiveness of teaching these activities to language learners, and the evidence for the effectiveness of the targeted strategies for learning L2 phonology. Chapter 3 focuses on descriptions of suprasegmental features of English, how they relate to segmentals, and the rationale for focusing on the targeted phonological content (message units, message unit boundaries, primary phrase stress, intonation, rhythm, word stress,

and multiword construction stress) when using the previously mentioned pronunciation learning strategies. Additional topics include the targeted content's roles in L2 intelligibility and instructional practices for improving learners' control of the targeted content. For each part, gaps in the relevant literature are discussed, thus pointing to the research questions proposed for the current study. Chapter 4 provides a description of the pronunciation course which was the context for this study. The purpose of the chapter is to document how the strategies and phonological content were presented to the participants. The study's methodology is explained in Chapter 5. Results of the study are presented in Chapter 6, and a discussion of key findings, limitations, pedagogical implications, and future research needs are presented in Chapter 7.

Chapter 2

Self-Monitoring and Language Learning Strategy Research

This chapter positions the current study within the LLS literature and discusses the following topics: definitions of the self-monitoring strategies investigated in this study; definitions of general language learner strategies and L2 pronunciation strategies; relating pronunciation self-monitoring strategies to the larger LLS literature; an exploration of what we know about teaching learners to use pronunciation self-monitoring strategies; research findings on the effectiveness of strategy-teaching techniques; evidence for use of self-monitoring to improve pronunciation; and gaps in the pronunciation strategy training literature.

Self-Monitoring Strategies

Before offering definitions and categories of LLSs and before discussing in greater detail how this study's L2 pronunciation self-monitoring strategies relate to strategies identified in the general LLS literature and the nascent pronunciation strategy field, I first provide descriptions of the current study's four self-monitoring strategies, critical listening, transcription, annotation (correcting a transcript), and rehearsing corrections aloud.

Critical listening. As I explain in more detail in a subsequent section, learners often are encouraged to listen to their own production in order to identify nontarget features (W. B. Dickerson, 1987, 1994; Morley, 1991). Yet how the learner goes about listening is quite important and little studied. Listening holistically may allow a learner to identify the most salient features, such as use of too many fillers and self-repairs or a lack of fluency. In this study**,** listening is meant to be more detailed and systematic, a critical process that provides an opportunity for learners to focus on their L2 production at the segment, syllable, word, phrase,

and discourse levels. When listening to recordings of their own speech, learners must replay and critically listen to short segments of the recordings multiple times, one message unit (MU) or perhaps even one word or syllable at a time. With each listening, the learner must focus on just one target feature, for example, placement of MU boundaries, presence or absence of primary phrase stress (PPS), phrase-final intonation, or stress of individual words. Learners are instructed to listen for specific pronunciation cues, for example, focusing on the pitch jump or drop on the syllable receiving PPS, the type of intonation following the PPS, and quicker, quieter de-accented syllables following the PPS, rather than try to listen more generally for PPS. When the learner is targeting specific cues such as these, rather than listening holistically, the repeated listening to these short segments of speech may help the learners notice what they are producing, identify nontarget features, and then identify changes to be made, thus facilitating bottom-up processing (as suggested by Izumi, 2003).

Transcription. This strategy cannot be separated from critical listening. In order for transcription to be effective, an L2 learner must also be able to attend to and identify the target pronunciation features in their speech. When transcribing speech, the learner writes down exactly what was spoken, including nontargetlike production of suprasegmentals, pauses, restarts, fillers, and repairs. The goal is to create an accurate written record of a speech sample, without yet moving ahead to the evaluation phase. The process of focusing on the phrase, word, syllable, and sound levels and transcribing (writing down what is heard) likely allows the learner to overcome limitations of working memory. Rather than focusing on complete utterances and attempting to keep information about many features in working memory, transcription allows the learner time to analyze and reflect on his or her output. Once a longer speech segment has been transcribed, the learner can review what was produced and identify sections that require further analysis, a

process that likely would be difficult if not impossible if the learner were to rely on listening alone.

Transcription may be completed by the learner or the teacher, and learners may transcribe their own speech or that of NSs. This study focuses on transcription by the learner of her or his own oral production rather than on learner transcription of NS models (e.g., Clennell, 1999). However, as I note in a later section, listening critically to the speech of others is an important part of the self-monitoring training used in this study. And because of the underlying goal of developing self-study skills, I also do not investigate transcription of students' speech by teachers[1] (Lynch, 2007).

Annotation. For this strategy, learners go one step further with the transcription described in the previous section. Learners review their own transcripts, looking for nontarget features, and annotate (mark) corrections directly on the transcript in a contrasting color. During this transcript correction process, learners refer to a checklist (Appendix A) to remind them of the pronunciation features they should monitor. Learners apply knowledge of the pronunciation rules learned in class (see Chapter 3; e.g., primary phrase stress goes on the last content word or function word in new information; all but the final item in a series typically has rising intonation). Reviewing the transcript is an evaluatory process that the learner completes "off-line", when they have the processing resources available to attend to form and its relationship to their intended meaning. Placement of MU boundaries, the type of intonation used, and PPS placement all are integral to their intended meaning. Thus attention to form and meaning remain

[1] See Chapter 7 for a discussion of the ways a teacher's transcriptions of learner speech can inform the teacher's understanding of learners' pronunciation strengths and weaknesses.

appropriately integrated, but because the meaning is already established by the learner, the learner can shift a larger portion of the focus to the correct, corresponding form.

Part of the rationale for using this self-correction process is that some evidence exists for the value of evaluating one's production after the fact as opposed to trying to plan accurate pronunciation prior to speaking. Foster and Skehan (1996) looked at accuracy in syntax, morphology, and word order and found that the "most accurate performance [was] produced by the less detailed planners" (p. 299). In his study of timing of self-monitoring, W. B. Dickerson (1987) found that learners were more accurate in correcting their own pronunciation after they monitored their own speech than when they had time to apply rules before speaking.

Rehearsing corrections aloud. The idea that practice is important in language learning and for L2 pronunciation in particular is not controversial. And within the LLS research literature, practice has been cited as an effective strategy and one of the more frequently used (Chamot & Kupper, 1989; Cohen et al., 1995; W. B. Dickerson, 1989; Sardegna, 2009). Thus the rehearsal phase used in the current study is meant to reflect the type of practice learners typically do when learning new L2 skills. After performing one of the self-monitoring tasks with a given speech excerpt (i.e., critical listening, listening + transcription, or listening + transcription + annotating corrections), learners orally produce each speech excerpt three times. I chose three rehearsals for pragmatic reasons: Going beyond three might result in fatigue or too great a cognitive load; and based on my own observations of students' attempts at self-monitoring, students tend to make increasingly more corrections following at least one or two rehearsals. The goal is to make their production as accurate as possible in terms of the target pronunciation features. During and following each rehearsal, learners monitor and evaluate their output, with

the goal of identifying modifications that are needed in subsequent rehearsals in order to make their oral production targetlike.

In the next section, I provide an overview of the general LLS literature and what is known about strategy-based instruction, what has been done so far regarding the investigation of pronunciation LLSs, and how the current study's targeted strategies relate to this literature.

Definitions of Language Learner Strategies

Researchers and theorists in the LLS literature (Cohen et al., 1995; Gu, 2007; Hassan et al., 2005; Macaro, 2006; McDonough, 1995; O'Malley & Chamot, 1990; Oxford, 1990) generally agree that LLSs are conscious (or likely start out that way), and are goal-directed actions carried out by learners themselves, in contrast to other factors that may impact language learning, such as teacher-directed instruction, "subconscious activity, language learning processes, skills, learning plans, and learning styles" (Macaro, p. 325). Similarly, strategies are seen as consisting of mental activity, and most, excluding Macaro, agree that strategies also can include observable learner behaviors. Strategies are used to aid and improve learning and are "aimed at maximizing results" (Gu, p. vii). And according to Macaro, LLSs "are the raw material without which L2 learning cannot take place" (p. 332).

Table 1 lists the most common definitions of LLSs of the past 35 years, from the earliest (Rubin, 1975) to more recent (Gu, 2007).

Table 1

Definitions of General Language Learner Strategies, in Chronological Order

Source	Definition
Rubin (1975)	"The techniques or devices which a learner may use to acquire knowledge" (p. 43).
Oxford (1990)	"Steps taken by students to enhance their own learning" (p. 1).
O'Malley & Chamot, (1990)	"The special thoughts or behaviors that individuals use to help them comprehend, learn, or retain new information" (p. 1).
Cohen, Weaver, & Li (1995)	"The steps or actions selected by learners to improve the *learning* of a foreign language, the *use* of a foreign language, or both" **(**p. 2).
Naiman, Frohlich, Stern, & Todesco (1996)	Strategies are conscious and "general, more or less deliberate approaches" used by language learners." Techniques are "observable forms of language learning behavior" (p. 4).
Hsiao & Oxford (2002)	"The L2 learner's tool kit for active, conscious, purposeful, and attentive learning" (p. 372).
Hassan et al. (2005)	"Any intervention which focuses on the strategies regularly to be adopted and deployed by learners in order to develop language proficiency, improve language task achievement or both" (p. 10).
Macaro (2006)	A strategy description should consist of "a goal, a situation, and a mental action" (p. 325).
Gu (2007)	"The learner's decision-making process and the behaviors involving learning decisions aimed at maximizing results" (p. vii).

LLS typologies vary most in how they define the size and scope of strategies. Some theorists define strategies as the overriding approach taken by a learner and then specify techniques or tactics that represent specific actional components, that is, the discrete behaviors or mental actions that comprise the strategy. For example, Peterson (2000) identifies one pronunciation strategy as "self-evaluating" and a corresponding tactic as "recording oneself to listen to one's pronunciation". Naiman, Frohlich, Stern, and Todesco (1996) distinguish between strategies (general approaches) and techniques (observable behavior). Others identify strategy hierarchies that do not attend to the variation in size, abstractness, or scope of the strategy. For

example, Oxford (1990) lists the strategies of self-monitoring (likely an internal, mental process) and "asking for correction" (an observable behavior involving other's input). Yet others seem to use strategy to refer both to the overriding approach and to the discrete behaviors that subsume them (e.g., O'Malley & Chamot, 1990, who identify the strategy of "directed attention," in which a learner "decides in advance to attend to a task and maintains attention during the task", p. 137). Macaro (2006) is perhaps the most clear in addressing the scope/size/abstractness conflation:

> A strategy's description should be effected at the lowest relevant level of articulation within the boundaries of conscious cognition….[I]t should not be possible to describe a strategy by referring to a number of relevant subordinate strategies….[S]trategies such as *rehearsing and memorizing* or *rereading texts* [O'Malley & Chamot, 1990] are in fact combinations of strategies. (p. 327)

The latter strategy, "rereading a text" (e.g., a text one has written with the goal of improving it) actually consists of strategies such as "does it sound right? does it look right? what are the mistakes I usually make? and so on" (p. 327). The author goes on to clarify that learners typically use strategies in clusters and that in fact they must be used in such combinations to be useful.

LLS theorists and researchers have generated a number of different strategy taxonomies. Most theorists based their lists on those originally developed by Rubin (1981) and Oxford (1990) and the most common overriding categories are metacognitive (thinking about and evaluating one's learning) and cognitive (working directly with the L2). O'Malley and Chamot (1990) suggest a third category: socioaffective strategies (which have to do with interacting with others and controlling one's affective states). Summaries of the most common strategy lists are presented in Appendix B.

Next, I describe what is known to date regarding LLSs used specifically for L2 pronunciation.

L2 Pronunciation Strategies

One researcher who initially addressed the role of pronunciation learning strategies was W. B. Dickerson (1984), who emphasized the importance of a particular strategy cluster, covert rehearsal (private practice), during which learners self-monitor and self-correct their pronunciation. Only in the past decade have researchers started to identify the full range of pronunciation learning strategies used by L2 learners (Derwing & Rossiter, 2002; Eckstein, 2007; Osburne, 2003; Pawlak, 2010; Peterson, 1997, 2000; Vitanova & Miller, 2002). Each has approached strategy identification in different ways and with varying degrees of analysis, thoroughness, and focus. For example, Dickerson developed a specific set of steps for covert rehearsal that systematically guided the learner through monitoring and correcting their L2 pronunciation (see Sardegna, 2009, for a complete description). Derwing and Rossiter were primarily concerned with how L2 speakers handled communication breakdowns due to pronunciation problems. The current study is not focused on learners' use of communication strategies, but rather on strategies used specifically for learning L2 pronunciation. Thus I will not discuss Derwing and Rossiter's findings. Osburne's categories directly targeted specific pronunciation features in a way that the other taxonomies have done more generally via strategies of noticing and self-correcting (Eckstein), self-monitoring (Eckstein; Vitanova & Miller), and self-evaluating (Peterson). Vitanova and Miller's findings were based on action research conducted in their language classrooms, with the goal of encouraging pronunciation self-awareness and strategy use. Eckstein, Osburne, Pawlak, and Peterson each used various observational, self-report, and oral protocols to uncover the pronunciation strategies learners were using. Eckstein, Pawlak, and Peterson were attempting to frame their strategy lists within pre-existing taxonomies or conceptual frameworks, whereas W. B. Dickerson, Osburne, and

Vitanova and Miller had other motivations or goals. My intent in this chapter is to provide brief descriptive information for each existing taxonomy and to discuss the extent to which the strategies targeted in the current study fit within these pronunciation-specific frameworks. In-depth critiques of how the best-known pronunciation strategy categories were developed and their respective strengths and weaknesses appear in Eckstein (2007) and Pawlak (2010).

W. B. Dickerson (1989, 2000). In Dickerson's (1989) text, *Stress in the Speech Stream,* the author places pronunciation instruction in a framework of helping learners develop independent learning strategies that include self-monitoring and self-correction. In class, students learn to use strategy clusters for applying pronunciation rules and then apply these strategies in a principled manner during private practice, or "covert rehearsal". The process of covert rehearsal includes the following six steps: find privacy to practice throughout the day, talk aloud in English, monitor performance for specific features, compare performance with models stored in memory, make changes in production to match the models, and practice changes out loud. These steps are repeated until the learner is satisfied that her or his production is accurate (W. B. Dickerson, 2000). Dickerson's work is the only source I found that developed LLSs specific to pronunciation, unlike the more recent group of writers, who have focused on strategy identification, rather than on strategy development.

In the current study, I used self-monitoring and self-correcting techniques largely based on to those described by W. B. Dickerson (1989) and L. D. Hahn and W. B. Dickerson (1999). I added three elements to the range of covert rehearsal techniques: learner use of self-recordings, self-transcription, and annotation of transcriptions (described in detail in a later section).

Peterson (1997, 2000). Peterson used diaries and interviews with 11 L2 learners of Spanish to develop a list of strategies for improving pronunciation. She then categorized them

into a list of 12 strategies (with associated tactics in parenthesis below) that were intended to be consistent with Oxford's (1990) strategy taxonomy:

1. Representing sounds in memory
2. Practicing naturalistically (concentrating intensely on pronunciation while speaking and while listening to the TL [target language]; talking aloud to oneself; concentrating intensely on pronunciation while listening to the TL; trying to avoid producing inappropriate native language sounds; imitating a NS)
3. Formally practicing with sounds (pronouncing a difficult word over and over; repeating aloud after tapes)
4. Analyzing the sound system (forming and using hypotheses about pronunciation rules)
5. Using proximal articulations
6. Finding out about target language pronunciation
7. Setting goals and objectives (deciding to focus one's learning on particular sounds; deciding to focus one's listening on particular sounds)
8. Planning for a language task
9. Self-evaluating (recording oneself to listen to one's pronunciation)
10. Using humor to lower anxiety
11. Asking for help
12. Cooperating with peers (pp. 25-26)

Vitanova and Miller (2002). These two language instructors asked students in three different pronunciation courses to reflect on their reasons for improving their pronunciation and to describe what they found "most helpful in improving pronunciation" (p. 2). Based on the students' responses, the authors identified two categories of metacognitive pronunciation strategies: self-monitoring and active listening to and mirroring (non-verbal modeling) of NSs.

Osburne (2003). Osburne elicited speech samples from 50 adult ESL learners from university and adult education contexts, then replayed excerpts for them and asked the participants to try to correct any errors. She then asked learners to describe what they were doing

to improve their pronunciation in each instance. After analyzing and categorizing learner responses, she identified eight categories:

1. Focusing on local articulatory gesture or single sounds
2. Focusing on sounds below the syllable-level
3. Focusing on individual syllables
4. Focusing on prosodic structures; monitoring global articulatory gestures
5. Focusing on paralanguage
6. Focusing on individual words
7. Focusing on memory or imitation (pp. 135-136)

Eckstein (2007). Eckstein offers a definition of pronunciation strategies based on Oxford (1990) and Peterson (2000): "Specific actions taken by the learner to make pronunciation learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations" (p. 12). The author administered his "strategic pronunciation learning scale" to 183 adult ESL learners and found "that strong pronunciation learners [generally] used pronunciation learning strategies more frequently than poorer learners" (p. v). The taxonomy he developed as a result of his research is categorized according to "four stages of pronunciation acquisition: input/practice, noticing/feedback, hypothesis forming, and hypothesis testing" (p. vi). The complete strategy list appears in Table 2.

Table 2

Pronunciation Strategy Taxonomy Developed by Eckstein (2007, p. 35)

Input
Intent listening
Focusing on articulatory gestures of others
Active listening
Eagerly listening to new sounds
Putting self in proximal points for hearing L2 pronunciation: TV, Movies, Radio, etc.
Representing sounds in memory
Focusing on individual syllables of words
Practice
Reading aloud
Practicing new sounds
Imitating and/or mimicry of native speakers
Practicing 'mock talk' or imitating L2 prosody using L1 words
Talking aloud/role-play
Memorizing the pronunciation of words
Helping facial muscles become accustomed to accommodating L2 pronunciation
Practicing different sounds, first in isolation and then in the context of words
Repeating after tapes in a language laboratory
Noticing
Noticing the intricate differences between L1 and L2 pronunciation
Focusing on suprasegmentals of language
Intent listening
Distinguishing errors among other speakers
Focusing on articulatory gestures of others
Listening carefully to errors made by native speakers to infer key sounds or structures
Acquiring a general knowledge of phonetics
Feedback
Self-monitoring
Focusing on suprasegmentals of own speech
Using phonetic symbols and transcriptions
Monitoring and eliminating negative interference
Active listening
Asking for help
Cooperating with peers
Hypothesis forming
Monitoring and eliminating negative interference
Self-correcting
Acquiring a general knowledge of phonetics
Doing special exercises for sounds not existing in the learner's native language
Finding out about the target language pronunciation

(continued)

Table 2 (continued)

Hypothesis testing
Repeating new words according to new hypotheses
Skipping difficult words
Rehearsing sounds
Using proximal articulations
Increasing or decreasing volume of speech
Using a slower rate of speech
Using clear speech
Lowering anxiety

Pawlak (2010). Pawlak's goal is to develop a valid instrument for identifying pronunciation strategy use. Though the author has not yet published the resulting survey tool, he provides useful definitions and cites others' research regarding the effectiveness of strategy use. For example, he states that pronunciation learning strategies are "deliberate actions and thoughts that are consciously employed, often in logical sequence, for learning and gaining greater control over the use of various aspects of pronunciation" (p. 191). And citing his own and others' work, he notes that such strategy use contributes to the development of declarative (explicit) and procedural (implicit) knowledge.

Relating Pronunciation Self-Monitoring Strategies to the LLS Literature

The targeted strategies (listening, transcription, annotating a transcript, rehearsing corrections aloud) most closely relate to some of the more broadly defined language learner strategies commonly identified in the general LLS literature, specifically self-monitoring, self-evaluation, and self-correction, and also include elements of deductive reasoning, prediction, revision, and practice/rehearsal (Macaro, 2006; Naiman, Frohlich, Stern, & Todesco, 1996; O'Malley & Chamot, 1990; Oxford, 1990; Rubin, 1981). Within the pronunciation strategy literature, this paper's targeted strategies are nearly identical to covert rehearsal as defined by W.

B. Dickerson (1984), but with the addition of transcription and annotation. Also, because learners completed the experiment in a lab setting, the tasks were not truly covert rehearsal. In Appendix C, I have listed for each targeted pronunciation strategy the relevant strategies from the general LLS and pronunciation strategies literature, organized by strategy type. How these specific LLSs relate to the current study's targeted strategies also is discussed next and is illustrated in Figure 1.

The specific activities listed in Figure 1 are not, in isolation, considered learner strategies. However, when used with the goals of monitoring, evaluating, and correcting one's L2 output, the discrete actions required to carry out each activity then function as language learner pronunciation strategies, that is, they are tools or goal-directed actions that L2 learners may select to improve their L2 pronunciation. These specific strategies function together to allow learners to self-monitor, self-evaluate, and self-correct their L2 pronunciation. When self-monitoring, learners must be able to listen critically to their own L2 production. Listening is used to help learners focus their attention on the target pronunciation features. When self-evaluating, learners compare their output to a model L2 target in working memory or apply L2 predictive rules and identify important discrepancies. Transcription is used to help learners see what they are doing and evaluate it. Annotation involves applying L2 knowledge and predicting corrected targets. When self-correcting, learners attempt to modify their production to more closely match the target, and rehearsing corrections aloud provides an opportunity for the learner to orally correct. After rehearsing orally, the learner again listens critically, evaluates, and self-corrects. It becomes a cycle of monitoring, evaluation, correction, which, over time, may reinforce the pairing of articulatory movements and the targetlike aural input (W. B. Dickerson, 1984). And most likely these three strategy clusters are not used separately, but rather the learner may use

the various strategy clusters at whatever point in the overall process that they are needed. They are presented here as separate processes for explanatory purposes. Also, when learners are taught to use these strategies, learning them in an organized fashion such as this facilitates the training process.

Macaro's (2006) definition requires that strategies represent "the lowest relevant level of articulation within the boundaries of conscious cognition" (p. 327) and he notes that strategies are used in clusters or organized into strategic plans. Using these constraints on strategy definition, I have identified the strategy clusters of self-monitoring, self-evaluation, and self-correction (*correct transcript* and *rehearse corrections aloud*). These clusters could be deployed when a learner's goal is to improve pronunciation, perhaps for a specific purpose, such as a teaching context (the strategic plan). The specific activities would be used by the learner as techniques (Naiman et al., 1996) or tactics (Peterson, 2000). Listening alone and in combination with transcription can be used for self-monitoring and self-evaluating; correcting a transcript and rehearsing corrections aloud would be used initially for self-correction, but would also require cycling through the three clusters (monitor, evaluate, correct), until the learner is satisfied with her or his production.

In Figure 1, I have indicated the specific strategies (goal-oriented actions that are mental or observable behaviors) that would be used to implement each strategy cluster. In this scenario, strategies are used in multiple situations for multiple purposes. In each cluster, it becomes necessary, for example, to use selective attention, focus on specific features, and apply prior knowledge about pronunciation. This feature of generalizability and transferability to other tasks is mentioned by Macaro (2006) as an important characteristic of LLSs.

One final point regarding the relationship of the strategy clusters to the LLS literature must be added. In a given teaching context, the definitions used for learner strategies need to be relevant to the learners and presented in a way that learners can immediately apply to their own language learning. Learners themselves may be less concerned with whether a strategy is defined only as mental action or an observable behavior, whether it represents the "lowest relevant level of articulation", whether or not is it conscious, and so on. Carefully defining and specifying such characteristics is critical for researching strategy use and effectiveness. We need to ensure that, in the LLS literature, we are using terms such as strategy, self-monitoring, or listening critically in a consistent manner so we can consistently define independent variables and interpret their associated outcomes. However, in the classroom, students benefit most from the use of strategy labels that are already familiar (such as self-monitoring, listening, correcting errors, transcribing, rehearsal) and for which the purpose and procedures are made clear during the course of instruction. In order to be consistent in usage of terms when I am describing the instruction offered in the current study's teaching context and discussing the study's results, throughout the remainder of this paper I will refer to the four targeted behaviors (listening, transcription, annotating corrections on a transcript, and rehearsing corrections aloud) as strategies.

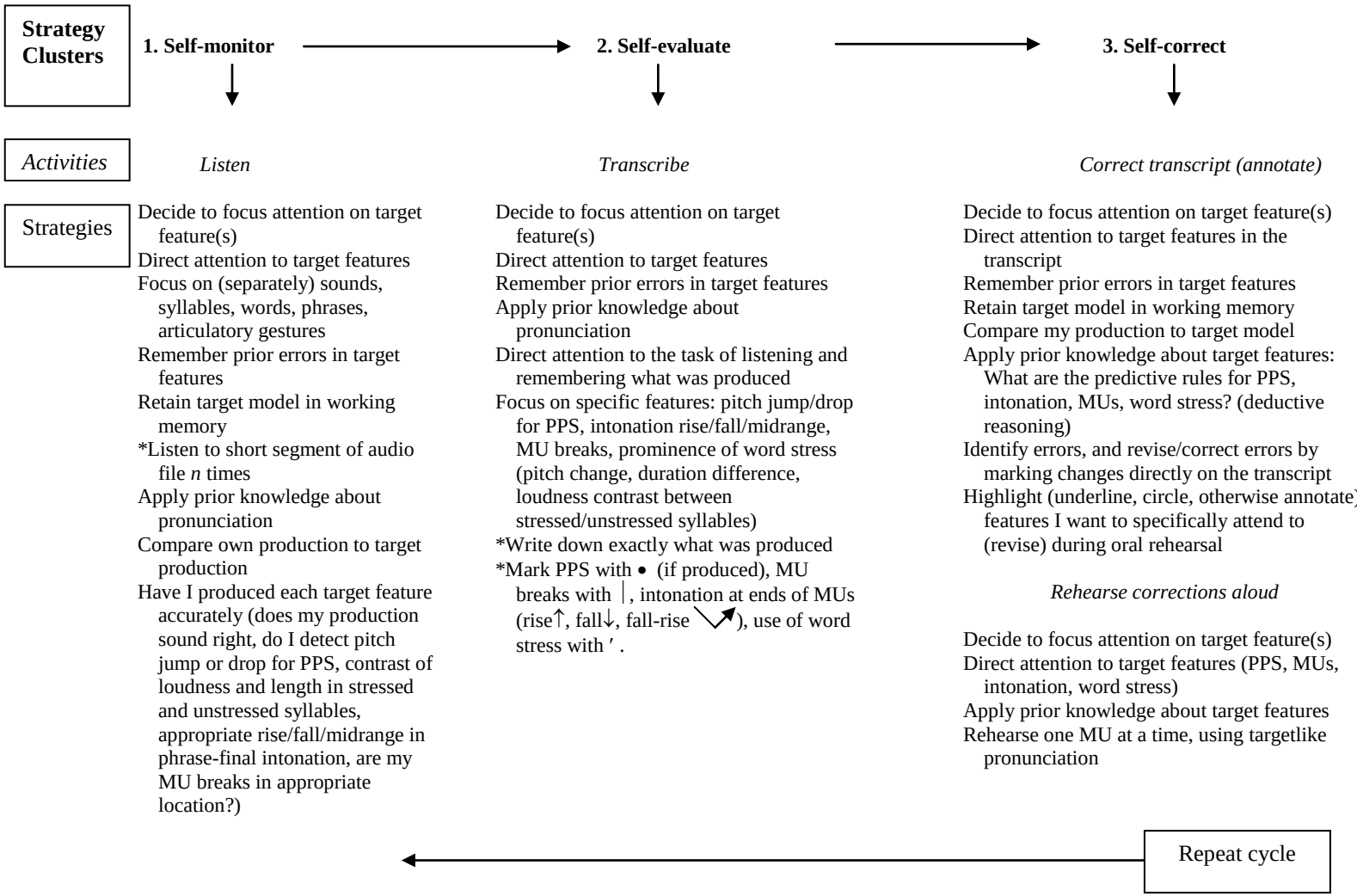


Figure 1. A sample strategic plan for improving L2 pronunciation for teaching. (* = activities not mentioned in LLS and pronunciation strategy literature)

To summarize, limited options exist for carrying out self-monitoring and self-evaluation for pronunciation: on-line listening or monitoring (listening critically to oneself while speaking), self-monitoring off-line (after speaking) by listening to a video or audio recording of one's speech, viewing an annotated transcription of one's production, and use of speech visualization software are the only techniques mentioned in the pedagogical literature (discussed in the next section). The strategies targeted in this study involve all of these options except for speech visualization. The rationale for excluding this option is to use only strategies that learners can easily implement, without requiring access to and training for use of specialized technology or resources not commonly available to most learners[2]. The selected strategies are worthy of study because they help learners focus their attention on target L2 features and give them a concrete approach to controlling their L2 monitoring and production.

Teaching Learners to Use Pronunciation Self-Monitoring Strategies

Learners likely are familiar with the concepts of listening to their own speech and monitoring, evaluating, rehearsing, and correcting it orally. However, they may not have had training for using these practices in a systematic and effective way. Additionally, students rarely use self-recordings and transcription as a means of monitoring, evaluating, and correcting their production outside the language classroom (Smith & Beckman, 2010). Though several writers describe how they have used transcription successfully in their classes (Acton, 1984; Blanche, 2004; Lynch, 2007; Mennim, 2003, 2007), none of the pronunciation strategy research (nor

[2] See upcoming section, "Developing L2 perception skills," for a description of studies that have investigated the effectiveness of learners' viewing of spectrograms of NS speech samples, in order to visualize pitch contours, intensity, and duration.

general LLS literature, for that matter) highlights the use of transcription as a specific strategy, technique, or tactic.

Learners often are most comfortable with receiving feedback from their instructors or those whom they perceive as more competent L2 speakers. Learners often report that they are unsure what to listen for and also are less confident in knowing how to consistently and accurately identify their errors and produce the target features accurately (e.g., Aufderhaar, 2004; M. G. Chang, 2006). Additionally, when learners first listen to recordings of their own speech, initial reactions are negative and highly critical (Y. M. Chen, 2008). In order to become competent and more neutral and objective in evaluating their own speech, L2 learners need to develop four prerequisite skills via classroom instruction and covert rehearsal: a sense of disinterest when listening to one's own speech and perception, prediction, and production of the targeted pronunciation features (W. B. Dickerson, 1989). In the following paragraphs, I describe techniques for developing these skills and then explain ways of training learners to use the targeted strategies.

Disinterested listening. In my own experience and as reported by Y. M. Chen (2008) and Cohen, Weaver, and Li (1995)**,** language learners require time to become neutral and objective when listening to their own speech recordings. Development of this objective stance can be facilitated in several ways: (a) Through explicit reassurance from the instructor that such feelings are normal but that objectivity is essential for effective self-monitoring; (b) through regular exposure to their own recordings (Y. M. Chen reported that her students became significantly more objective during their second self-assessment task); (c) and through demonstrations that oral English is "messier" than written English and that even NSs' oral production contains disfluencies, fillers, and self-repair. In an activity illustrating this third item,

I present students with a transcript of my own speech from a previous ESL course for which I was the instructor. The transcript is all lower case, contains all the fillers, restarts, and features typical of spoken discourse. Students predict where endings of utterances are most likely to occur and are asked to give their evaluation of the speaker's English proficiency. Then learners listen to the recording, are very interested to discover that the speaker is a NS of English, and subsequently discuss what makes spoken English different from written forms. These and other tasks can be used to help learners become disinterested listeners.

Developing L2 perception skills. Learners must become familiar with how the target features sound in NS speech. In the current study, perception training involved a focus on NS models, which included the instructor and audio files from websites such as *Scientific American* and *How Stuff Works*, and the learner's own speech (items a-d in the following list). Perception training may be accomplished in a number of ways: (a) Through in-class descriptions and live and recorded NS models of the most salient features of the targets. For example, the most salient features of PPS are pitch change, increased loudness, and lengthened vowels on the stressed syllable, contrasted with quicker and quieter syllables in any words that follow within the same MU. Learners listen to recordings or examples spoken by the instructor, who repeatedly highlights the key features for the learners, replaying or repeating the models as often as needed. When appropriate, the equivalent of minimal pairs can be used to help contrast and pinpoint the relevant pronunciation cues. (b) Homework and classroom activities reinforce perception of the features, through learners' listening to audio files of NS speech and identifying the targets on a transcript of the oral text. Instructors can monitor learners' perceptual accuracy and offer direct corrective feedback through written notes on or recorded audio responses to homework. (c) Instructors also can provide feedback and work with learners individually during one-on-one

conferences outside of class. (d) Learners may record their own speech and practice listening for the presence or absence of the target features (Couper, 2003). (e) Students may also work in pairs and offer peer feedback on each other's production of the target features (Lynch, 2007) or guide each other through self-monitoring steps or "queries" (W. B. Dickerson, 1989). (f) Some instructors and researchers have investigated the effectiveness of viewing spectrograms of NS speech samples, in order to visualize pitch contours, intensity, and duration (Anderson-Hsieh, 1992, 1994; deBot, 1983; Seferolug, 2005). For some learners, being able to "see" these features helps them to better perceive the targets in their own and others' speech. As noted previously, in this study visualization is not used, due to the focus on techniques that learners can use independently, without the use of resources that require specialized linguistic knowledge or additional training to use. Learners with access to inexpensive sound recorders, smart phones, or to computers with built-in microphones and speakers should be able to use the self-monitoring strategies investigated in this study.

Overall, the pedagogical and research literature supports the use of a combination of instructional and self-study practices for reinforcing the development of perception skills for the targeted pronunciation features, including in-class listening to NS target models, feedback from an instructor and peers, self-evaluation, viewing visual output of the speech signal, and homework listening activities.

Developing L2 predictive skills. Next, learners develop declarative knowledge about the target features. The instructor provides explanations of how the target features are used in the L2: placement of MU boundaries (signaled by PPS and the following intonation and pauses following logical thought groups or grammatically complete phrases); use of PPS and word stress (use of increased vowel duration, pitch change, and increased intensity on the most

prominent syllable in the phrase or word; for PPS, prominence usually is contained in the word that is being highlighted due to new/old information distinctions or contrastive stress); intonation patterns (use of rising, falling, or fall-rise, following PPS) for statements, non-final phrases, series, and questions; and rhythm (use of linking of adjacent words in a MU; alternating stress, in which content words receive heavy stress and function words are unstressed, stressed syllables contain full vowels, and unstressed vowels are reduced). Classroom and homework activities include tasks that focus on rule-based prediction of where these features can occur. Students are given transcripts of dialogues or monologues and predict location of PPS, MU boundaries, intonation contours, rhythm, and word stress. After predictions are complete, learners listen to a NS model and compare their predictions to what was produced and discuss reasons for any discrepancies.

Developing L2 production skills. As the students are developing perceptual and predictive skills, they also practice producing the targeted content. Students complete in-class activities that help them learn what it feels and sounds like to produce lengthened vowels (through stretching of rubber bands to simulate lengthening and shortening of syllables; use of nonsense syllables, such as TAAA-TA, to keep the focus on lengthening and avoid the distraction of worrying about accurately producing segments: Chela-Flores, 2001); to get the feel of alternating rhythm (tapping hands or feet to the rhythm; using "build-ups", in which they say phrases that at first contain only content words and then function words are added, while the original rhythm is maintained, e.g., L. D. Hahn & W. B. Dickerson, 1999). Shadowing (repeating NS speech immediately following the speaker, word-for-word, imitating suprasegmental features), mirroring, and imitation (listening to an entire utterance and repeating it as it was spoken) are useful for students who are initially gaining control over these features. Hand

gestures and other physical movement may be used to signal and reinforce direction of intonation and stress placement while speaking.

Students also complete homework audio recordings, using their own computers or those available in campus labs. These recordings are uploaded to a web-based course management site such as Moodle (or could be submitted on audio cassettes or via cloud storage, e.g., Dropbox) to be evaluated by the instructor. In these recordings students practice producing the target features. As they gain more control, then learners work on rehearsing speech in meaningful contexts (such as minilectures or micro-teaching), identifying and planning where they should be using each target feature until use of the features gradually become more automatic.

Training learners to use the self-monitoring strategies. As disinterested listening, perception, prediction, and production skills develop, learners also are trained to use self-monitoring strategies. Specific steps for listening and transcribing are given: Listen repeatedly to MU-length segments of their own or others' recorded speech; listen for one targeted feature at a time; listen for specific features such as pitch change, loudness, vowel or syllable duration; not moving on to the next target or MU until one has accurately identified or transcribed what was spoken. Similarly, when correcting the transcription, learners attend to one target at time, reading through the transcript repeatedly until it has been reviewed thoroughly for each feature. Learners recall the prediction rules for each target and identify how each might be used appropriately and what needs to be corrected in the transcription. While listening, transcribing, or making oral corrections, learners attend to previous errors and double-check that they have not made the same error. Learners also are trained to rehearse in a systematic way, by focusing on one MU at a time and evaluating their production accuracy after they speak. During classroom instruction, students' nontargetlike production is not corrected directly. Instead students are guided through

the process of critiquing and correcting their errors through a use of questions or queries, as described by W. B. Dickerson (1984). For example, if PPS is misplaced, the student is asked to identify new and old information and find the last content or function word in the new information.

Though the ultimate goal is for learners to be able to use the self-monitoring strategies independently, these activities can be practiced by having students work with classmates to transcribe and evaluate each other's speech or a dialogue that was created collaboratively (Lynch, 2007). During the training phase, instructors can transcribe for students, provide feedback on students' transcriptions, and record students' transcriptions as a model of how the segment should sound. Additionally, individual conferences with students may be used to give additional feedback and training.

In the current study's teaching context, students had three opportunities during the semester for larger-scale contextualized practice of the targeted strategies and targeted pronunciation content. They produced three 5-minute "minilectures" (simulated teaching). These minilectures, which were audio recorded, occurred at the beginning, middle, and end of the semester and were presented in the classroom to their classmates. Following each presentation, participants completed homework assignments in which they listened to their own recordings and transcribed a 1-minute segment. They were instructed to broadly transcribe word-by-word, exactly what they said, including pauses, restarts, repairs, and errors (transcribed in standard English; they were not trained to use phonetic transcription). They focused on their use of the content studied in class. After completing the second and third transcriptions, students marked corrections on the transcripts, and then recorded the material again, trying to correct their errors. After all but the final minilecture, the teacher checked the accuracy of the transcriptions and

provided corrective feedback to the students. Also following each minilecture, students completed written self-reflections on their performance in the minilecture, commenting on overall performance and their abilities to perceive and produce the targeted suprasegmental features.

Summary. Classroom instruction and homework focus on the development of several key skills: disinterested listening, perception, prediction, and production of suprasegmental features in English and critical listening, transcription, annotation, and rehearsal of corrections. The procedures described parallel those explained by O'Malley and Chamot (1990, p. 158) for effective strategy-based instruction, and by W. B. Dickerson (1989) for pronunciation acquisition and use of covert rehearsal, which I have adapted here to relate to the targeted content and strategies:

- Students overcome self-consciousness when listening to their own speech and learn to be disinterested but critical listeners
- Students become aware of the target pronunciation features and value of pronunciation strategies such as self-monitoring, self-evaluation, self-correction
- Students are exposed via classroom activities to models of the target pronunciation features and strategies
- Students practice using the pronunciation features and strategies in class, in homework assignments, and in contextualized activities such as minilectures and subsequent transcription and correction
- Students learn to evaluate their production of targeted content and strategy use through written self-reflections
- Students learn how the strategies can be used for other tasks such as monitoring other pronunciation and oral English skills

Effectiveness of General LLS Training

In general, LLS training appears to aid language learning. Hassan et al. (2005) conducted a review of 38 LLS training studies completed between 1981 and 2002. Participants ranged from high school through adult and 24 of the studies were conducted in ESL (14) and EFL (10) settings. The others involved foreign language instruction in seven other commonly taught languages. Intervention length among the studies ranged from less than 2 hours up to 52 weeks. They found that simply using LLS awareness-raising training (one 50-minute session) did not result in increased strategy use in university foreign language students (Feyten, Flaitz, & LaRocca, 1999), indicating more extensive training is necessary. In general, extended LLS training for reading and writing are particularly effective for boosting learners' L2 performance. The three studies that looked at LLS training for improving oral skills similarly found positive results, including increased oral accuracy and better discussions. None of the reported studies investigated training learners to use pronunciation strategies.

More recent studies add to Hassan et al.'s (2005) findings about general LLS use. First, over the course of a 10-week strategy training period, university EFL learners were able to improve their oral skill self-assessments to match the level of the teacher's assessment (Y. M. Chen, 2008). Chamot and Kupper (1989) completed a longitudinal study of high school foreign language learners in which students were explicitly taught LLSs that were linked to specific tasks such as listening, speaking, and reading. They found that effective use of LLSs resulted in better language learning and they found self-monitoring to be one of the core strategies that was particularly useful for learners.

Cohen et al. (1995) found that 32 intermediate foreign language learners at a U.S. university benefited from 10 weeks of strategy-based instruction for at least one of three oral

tasks. Learners were explicitly trained to use LLSs for improving speaking proficiency and strategy use was embedded in classroom activities. Students were free to choose the strategies presented by their instructors (experimental group) or those that they already were familiar with (comparison group). Strategies that involved planning ahead, self-monitoring speech, and self-reflection on performance seemed useful in helping experimental group students improve grammar in a description task. Learners who reported paying attention to pronunciation were rated as improving in grammaticality of production and vocabulary usage. Learners who practiced the pronunciation of specific words before retelling a story were better at identifying key elements of the story during the retelling and were rated as more confident and more grammatical. Both the experimental and comparison groups benefited from the meta-cognitive process of reviewing performance and thinking about what to do differently in the future. Though overall the experimental group was more successful, learners in the comparison group actively used strategies to improve their performance. Several key differences exist between this study and my own: The study was not focused on pronunciation and related strategies. Learners were allowed to choose the strategies that they wanted to use for a given task. Also, information about strategy use was obtained from strategy checklists that participants completed following each of three oral tasks. Self-reported strategy use was then correlated with speaking scores. In the current study, my goal was to focus on a specific set of strategies to identify their effectiveness for a specific group of learners. Rather than relying on learners to accurately self-report their strategy use, I wanted to observe learners using the target strategies. And rather than using correlation to measure strategy effectiveness to a more general proficiency measure, I measured changes in specific pronunciation targets.

Perhaps the most useful insights coming from the general LLS literature have to do with developing an instructional framework for LLS instruction. Chamot and Kupper (1989) suggest the following based on their longitudinal study: (a) identify learners' current strategies (e.g., using retrospective and think-aloud methods, in small groups in the classroom); (b) assess their strategy needs (in relation to course objectives and the demands of the learning tasks students will perform); (c) plan the strategy instruction; (d) teach strategies explicitly and tie strategy instruction to specific language skills (the teacher should explain why the strategy is useful and model the strategy); (e) provide extensive strategy practice opportunities; (f) help learners to evaluate their strategy use; and (g) help learners to understand how to transfer strategies to other tasks. Items c through g were components of the strategy instruction in the current study.

What is Known About Pronunciation Strategy Training?

Many writers on pronunciation pedagogy emphasize the importance of activities similar to the self-monitoring strategies targeted in this study. Self-monitoring is perhaps the most frequently mentioned and rarely defined (Arteaga, 2000; Celce-Murcia et al., 2010; Firth, 1992; Kenworthy, 1987; Morley, 1991; Scarcella & Oxford, 1994). W. B. Dickerson (1989) and Sardegna (2009), however, offer in-depth descriptions of self-monitoring and self-correction (components of covert rehearsal). Transcription is used by Clennell (1999) to draw learners' attention to pragmatic aspects of spoken L2 discourse. Very few studies have checked the effectiveness of techniques for teaching learners to use these strategies. Of the seven studies I could find, only a few used adequate controls to allow reasonable conclusions regarding whether the techniques were learned. Next I briefly describe the seven studies in which learners were taught to use some of the self-monitoring strategies. Study summaries appear in Table 3.

Teaching learners to use listening, transcription, and oral rehearsal. Acton (1984) reports on instructional techniques used with adult professionals who needed to improve their intelligibility in English in order to be successful in the workplace. Learners listened to their self-recording; used oral rehearsal and transcribed and evaluated their interactions with NSs. Learners were trained on several techniques designed to develop learners' sense of suprasegmentals: "post-hoc monitoring", in which they "scan[ned their own] speech after the fact" (pp. 76-77); "kinesthetic monitoring", in which learners monitored for the "correct 'feel' of the target sound or process" (p. 77); tracking (repeating a NS's speech word-for-word); and mirroring. Acton cites evidence for improvement based on independent judges' rating of pre- and post-instruction speech samples and learners' self-report. However this evidence must be interpreted with caution, because (a) his goal was not to isolate the effects of the different instructional techniques, (b) he provided no descriptive or inferential statistics indicating the significance of the pre- and post-instruction gains, and (c) he included no comparison group following different pedagogical techniques.

Training on covert rehearsal and critical listening. W. B. Dickerson (1987; 1994) emphasizes the role of covert rehearsal, during which learners practice target L2 features, listen critically to self-monitor, and then correct their pronunciation. The 33 participants (equal groups of Chinese, Japanese, and Korean L1 speakers) in his 1987 study on self-monitoring and application of word stress and vowel quality rules were able to correct their pronunciation after they first produced an utterance and checked it for accuracy. As with the other studies, the goal of this one was not to compare pedagogical techniques. The goal instead was to evaluate optimal timing of rule use (in this case, after an utterance is produced), a finding that has highly relevant pedagogical implications.

A more recent study of the effectiveness of pronunciation strategy use is a classroom-based longitudinal study by Sardegna (2009), which took place in a teaching context very similar to my own. Sardegna evaluated the effectiveness of covert rehearsal strategy training for improving 39 university ESL learners' accuracy for primary phrase stress, construction stress, and word stress. She assessed learners' target accuracy at the beginning of the semester (time 1 [T1]) and found that learners' use of covert rehearsal strategies resulted in significant increases in accuracy on all three pronunciation targets by the end of the semester (time 2 [T2]). She noted persistence of learning 5 to 24 months post-instruction (time 3 [T3]) and an accuracy decrease at time 3 that remained above time 2 levels. Assessment at time 4 (9 months after time 3) indicated a plateau, with accuracy still significantly higher at time 4 than at time 2. The results for T1 to T2 (one semester of training/instruction) were as follows:

Feature	T1	T2	Difference
Primary phrase stress	55.97%	80.00%	24.03%*
Construction stress	43.72%	63.33%	19.61%*
Word stress	61.82%	82.95%	20.23%*

*Significant at the $p = .05$ level.

The common strategy components in both Sardegna's and my studies include critical listening, self-monitoring, rehearsal, and self-correction. A key difference is that Sardegna used a sentence-reading task to elicit the target features both pre- and post-instruction and thus production of the target features in spontaneous speech was not measured.

Listening training. Couper (2003) asked 15 post-intermediate ESL learners to record themselves imitating a NS. They then listened to their recordings and compared their production to the model's. He does not indicate whether the learners were to listen in a particular way (e.g., by focusing on one feature at a time) or a specific number of times. Couper counted errors produced pre- and post-instruction on sentence-reading and spontaneous production tasks and

found that errors decreased. He did not use a control group or provide repeated measures and did not test for statistical significance of the decrease in errors.

Training on transcription and annotation. Mennim (2003) described three Japanese university EFL students' use of self-transcription and annotation. First, students transcribed a 5-minute segment of their rehearsed 20-minute presentation. The teacher reviewed the transcripts, marked errors the students missed, and returned the marked-up transcriptions to the students. Students performed their lectures again, 1 week after receiving their teacher-corrected transcriptions. Teacher feedback focused on errors in article use, prepositions, passive structure, pronunciation (segments and word-level pronunciation), and elaboration of content. His study did not include a comparison group or repeated measures and he provided limited descriptive data and no statistical analysis regarding significance of improvements. His subsequent study (2007) was descriptive in nature. He tracked learners' use of a targeted form (noncount nouns) over a 9-month period, documenting some improvement in use during that time, which he attributed to the use of transcription. Mennim did not provide a detailed description of how the learners were trained to use transcription and self-correction.

Lynch (2007) compared the effectiveness of student-initiated (SI) transcription to teacher-initiated transcription (TI). In the former (SI), pairs of learners transcribed recordings of their planned and rehearsed role plays and worked together to make corrections to the transcript before submitting them to the teacher for further correction and reformulation. In the TI context, the teacher listened to recordings created by the learners and selected portions with errors to transcribe so the students could review and correct their errors. In both interventions, learners were able to correct the majority of highlighted errors when they performed the role plays. After a third performance, the SI group corrected 64% of their errors at time 1 and the TI group, 47%.

Though the focus was on grammar and vocabulary, his studies offer some insights into how transcription and correcting a transcript can be used for self-monitoring of L2 pronunciation. Lynch does not describe how learners were trained nor the amount of training provided.

Table 3

Summary of Studies That Trained Learners to Use Listening, Transcription, Annotation, and Oral Rehearsal

Study	Controls/comparison groups?	Findings
Acton (1984)	No. Classroom-based.	Learners' intelligibility improved after use of listening, transcription, oral rehearsal.
W. B. Dickerson (1987)	Yes.	Use of critical listening, self-correction, and oral rehearsal resulted in more accurate production.
Couper (2003)	No.	Listening to and evaluating one's production was associated with decreased errors in production on targeted features.
Mennim (2003, 2007)	No. Classroom-based.	Use of self-transcription, annotation, and self-correction was associated with improvement in pronunciation (2003) and use of noncount nouns (2007).
Lynch (2007)	Two quasi-experimental groups.	Use of student- and teacher-initiated transcription was associated with increased accuracy in grammar and vocabulary use, with the former showing greatest improvement.
Sardegna (2009)	Classroom-based. Learners functioned as their own controls.	Learners improved in the production of word, phrase, and construction stress following 16 weeks of covert rehearsal (critical listening, self-correction, and oral rehearsal) and pronunciation instruction.

Summary. Very few studies have looked at techniques for teaching learners to use self-monitoring strategies. I found three studies that looked at covert rehearsal (Acton, 1984; W. B. Dickerson, 1987; Sardegna, 2009); four for transcription (Acton, 1994; Lynch, 2007; Mennim,

2003, 2007); two for annotation (Lynch, 2007; Mennim, 2003); and four for critical listening to one's own speech (Acton, 1984; Couper, 2003; W. B. Dickerson, 1987; Sardegna, 2009). I was unable to find studies that specifically investigated training learners on the use of rehearsal of oral corrections, though it frequently is mentioned in pronunciation texts and in the LLS literature as a valuable strategy. Two studies (Lynch, 2007; Mennim, 2007) did not study effects on pronunciation, though their findings likely could be generalized to other features of oral English, in that they are using the techniques to draw learners' attention to aspects of their oral production. None of the studies specifically investigated how the strategies were taught. Only Lynch compared teaching techniques for helping students learn to use the activities. Except for W. B. Dickerson's (1987) findings on use of explicit rules, one can only infer that any improvement in oral production noted in these studies may indicate that the strategies were indeed learned and used as expected. Because of the lack of documentation or control in the bulk of the studies, one cannot claim that learners were using the techniques successfully.

Clearly more research needs to be done to better understand the effectiveness of teaching learners to use self-monitoring strategies.

Evidence for the Effectiveness of Self-Monitoring for Improving L2 Pronunciation

The seven previously discussed studies were used as examples of how learners have been taught to use self-monitoring strategies. As just mentioned, these researchers were interested in detecting improvements in learner L2 production due to the use of the strategies. Except for W. B. Dickerson (1989) and Sardegna (2009), the other researchers were not specifically testing whether the pedagogical techniques for training learners were effective, but were describing how the techniques worked for their learners. These same seven studies can also be used in discussion

of how effective the strategies themselves are for improving pronunciation. As mentioned already, these studies were not performed to compare and contrast use of the strategies nor were they used to determine each strategy's impact on the specific pronunciation targets, other than in very general terms. In each study, the results were positive, though due to weaknesses in several of the studies' designs, strong claims cannot be made regarding the effectiveness of using the targeted strategies for improving the suprasegmental features targeted in the current study.

I found only two studies that looked at the role of listening in improving L2 pronunciation (Acton, 1984; Couper, 2003). Both found positive results following use of their instructional practices, but neither study could offer definitive claims about the effectiveness of listening.

Instructors have used transcription and correcting of the transcript (Acton, 1984; Lynch, 2007; Mennim, 2003, 2007) in their classes and have used it to focus on pronunciation improvement, but generally the intent is for the instructor to give feedback, rather than to train learners to self-monitor and self-correct. Only the Lynch (2007) study provided a comparison of strategy types: (a) transcriptions created by and corrected by the learner with (b) transcriptions created by and corrected by the teacher *for* the learner. Both types of transcription were effective, though learners corrected more errors in post-task production on the student-initiated transcriptions. This finding supports the value of student-initiated transcription, and is relevant for the current study. The process of having learners work with an uncorrected transcript provided by the teacher has not been investigated.

Some research has looked at the effect of task repetition or recycling using communicative tasks (Bygate, 2001; Lynch & Maclean, 2001), concepts which are similar to what most instructors consider rehearsal. Both studies found improved accuracy in pronunciation

and other language features. Lynch and Maclean found that when adult L2 speakers gave the same poster talk repeatedly to different sets of listeners, they automatically made self-corrections to pronunciation in subsequent talks. Lynch (2007) describes the context of this research and explains,

> it has been argued that task recycling of this sort can allow learners to exploit their familiarity, gained during first performance, with the content and task demands, and with the process of formulating the desired meanings, so that they are able to devote more attention to getting the language right. (p. 312)

A study by Trofimovich and Gatbonton (2006) attempted to address the issue of rehearsal in a way that may have some relevance to the current paper. They conducted a series of word-priming experiments with 40 L2 learners of Spanish studying at a U.S. university to test the hypothesis that repetition and focus on form lead to improved perception (participants' reaction times decrease as they become more familiar with targets), which in turn leads to improved production. Their hypothesis was that familiarity with the L2 form frees up processing resources so the learner can attend to the form of input and output, rather than also needing to focus on meaning at the same time. They attempted to integrate the idea of repetition, creation of familiarity with L2 features, and use of communicative activities for teaching L2 intonation, though they do not provide empirical evidence for the effectiveness of their instructional practices.

As noted earlier, W. B. Dickerson (1987) investigated how the use of rules to monitor one's output leads to increased accuracy. This is one facet of his concept of covert rehearsal. Such rehearsal is not simply task repetition, but involves focused monitoring of one's output, use of rules to evaluate output, and then self-correction based on what was discovered from the monitoring.

In summary, few very general references have been made in the research and pedagogy literature regarding the use of self-monitoring as a general LLS and the implication has been that such use is related to learner pronunciation success. Reiss (1985; cited in Peterson, 1997) found that monitoring was the most commonly used of Rubin's (1981) strategies by the 98 foreign language students in her strategy study. In Breun's (2001a) study of 100 second-year Irish university learners of German as a foreign language, the researcher found that, among successful learners, one of the top 10 strategies used was "I try to notice my language errors and find out reasons for them" (p. 221).

Though language instructors and researchers on pedagogical techniques for pronunciation improvement have been interested in self-monitoring strategies of listening, transcription, annotation, and rehearsal, none has investigated all four in a systematic way, in order to determine their impact on learners' production of the targeted content. The studies described here indicate that these strategies can be useful and likely are effective for improving L2 suprasegmentals, but more empirical and appropriately controlled study is needed to tell us how the targeted strategies function in various combinations and if they have varying impact on the eight targeted pronunciation features.

Relationships Between Proficiency and Strategy Use

Proficiency measures. Knowledge of a group of learners' oral L2 proficiency helps language instructors set priorities for instruction, informs them regarding what to expect from their learners during a term of instruction, and guides the development of instructional materials. Oral L2 proficiency, or the level of oral competence achieved in the target language, may be measured in numerous ways. Holistic ratings of read and/or spontaneous speech are very

common, whether the test is a standardized one, such as the ACTFL Oral Proficiency Interview (Breiner-Sanders, Lowe, Miles, & Swender, 1999), Test of Spoken English (TSE), and more recently the Test of English as a Foreign Language Internet-based Test (TOEFL iBT), or one that is created locally for a specific institutional or teaching context. Holistic ratings, however, are not designed to specifically identify which aspects of oral production are accurate or inaccurate, particularly when ratings are based on multiple categories of competence, such as discourse, coherence, and sociolinguistic and linguistic accuracy. Even when evaluating L2 pronunciation, numerous factors often are evaluated holistically: accentedness, intelligibility, comprehensibility, fluency, and accuracy (Anderson-Hsieh, Johnson, & Koehler, 1992). Another approach is to elicit speech samples and evaluate accuracy on specific targets. For example, Derwing and Munro (1997) had raters transcribe the recorded speech of L2 learners, all of whom had read the same paragraph. Any word that was unintelligible was counted as an error. Such counts were used to arrive at a proficiency score. Even more detailed tests exist, such as the English Placement Test used at the University of Illinois at Urbana-Champaign (UIUC). This oral interview focuses on the same pronunciation features that are targeted in the university's remedial pronunciation courses. Such a test counts errors on specific pronunciation targets (sounds; word, phrase, and construction stress; linking; vowel reduction; intonation). Though the test is used for placement, it still gives a sense of a learner's pronunciation proficiency on separate pronunciation targets. Sardegna (2009) used this and similar types of tests for diagnostic and achievement purposes. In the current study, my definition of proficiency also is focused on this detailed level. The ESL course which was the context of the current study focused on instruction on specific English suprasegmentals, and the self-monitoring strategies were intended to help learners track their use of these same features. Thus, for this current study, I used counts

of targetlike tokens for each of the eight suprasegmental features from the learners' first minilecture as a baseline proficiency score, representing learners' accuracy in spontaneous speech. And though such speech samples generally do not fully represent a learner's L2 pronunciation competency, using more controlled elicitation of some kind would not offer measures for establishing proficiency that would also be equivalent to the experimental tasks.

To summarize, in this study, I am interested in the level of accuracy that learners have obtained on specific pronunciation targets. I am equating such accuracy with pronunciation proficiency. Accuracy measures for target pronunciation features are a direct measure of the effectiveness of the classroom instruction, which focused on specific suprasegmental targets. Holistic tests provide useful measures of overall intelligibility and comprehensibility, but do not identify learners' success in producing specific suprasegmental features.

Proficiency and general LLS use. As noted in previous sections, L2 researchers have been interested in identifying the ways that successful and less successful L2 learners are different, with the assumption that LLS training could result in less successful learners achieving greater proficiency. When looking at the research findings on LLS use, numerous studies indeed demonstrate a positive relationship between LLS use and L2 proficiency. That is, the more successful language learners tend to use more and a greater variety of LLSs. However, whether strategy use leads to higher proficiency or whether higher proficiency learners are better equipped or more likely to use LLSs effectively remains unclear. A review of the LLS literature indicates that the direction of the relationship has not been firmly established, though several useful observations can be made.

Most research has focused on LLS use focused on L2 writing, vocabulary development, grammar, testing, and overall oral skills. Very little has been written about pronunciation

strategies and their use by learners at different proficiency levels. The following discussion starts with a summary of the general LLS and proficiency literature, and then moves to what is known about pronunciation strategy use in relation to L2 proficiency.

General LLSs. Most studies on LLS use and L2 proficiency are correlational, comparing performance on one or more tests of L2 proficiency to learners' self-report of strategy use. This is a very reasonable starting point for determining if a relationship actually exists. The most common proficiency measures include standardized foreign language tests, course grades, oral interviews, cloze tests, picture description tasks, and jigsaw tasks use for speech elicitation. Self-report data often are elicited via strategy inventory surveys (such as the SILL, Strategy Inventory for Language Learning, Oxford, 1990) or through diaries, think-aloud protocols, or stimulated recall (White, Schramm, & Chamot, 2007). Generalizations from LLS studies are limited by the fact that learners are not observed using specific strategies, nor, in most studies, have learners received strategy instruction. Self-report data do not give a complete view of a learner's actual LLS use, effectiveness and appropriateness of strategies choices, and actual frequency of use. From these studies, we cannot know when or how learners decided to use a particularly strategy. What is missing in the research literature is an exploration of the quality and flexibility of strategy use by learners at various proficiency levels. Next I describe what is known from the correlational studies, and then I discuss the few studies that have observed learners' instructed strategy use in a classroom context.

Correlational and observational studies. Findings from studies relating self-reports of strategy use to L2 proficiency generally note a positive relationship between frequency, type, and quantity of strategy use and higher L2 proficiency levels. Following are key findings that are useful in guiding future research and pedagogy:

1. Higher proficiency learners use more strategies than lower learners and strategy use increases over time (Breun, 2001b; Chamot and Kupper, 1989; Lai, 2009; Oxford & Nyikos, 1989; Takeuchi, Griffiths, & Coyle, 2007).

2. Intermediate learners often use more meta-cognitive strategies than beginners, but both groups overall use cognitive strategies more than meta-cognitive ones. Repetition was the most common strategy for both groups and self-monitoring was used somewhat more by the intermediate group than by the beginners. These findings suggest that learners may benefit from explicit instruction on meta-cognitive strategies (Chamot & Kupper, 1989; O'Malley, Chamot, Stewner-Manzanares, Kupper, & Russo, 1985).

3. Use of self-initiation, selective attention, and oral repetition are positive predictors of general L2 proficiency (Gu and Johnson, 1996).

4. Effects of strategy use depend on the type of task and the combination of other strategies used. "Strategic competence exerts a causal effect on performance" (Purpura, 1997, p. 311).

5. Higher proficiency students use social strategies more than lower students (S. J. Chang, 1990; Chesterfield & Chesterfield, 1985).

6. Less successful learners use many of the same strategies as successful learners, but do not apply strategies appropriately and lack the meta-cognitive skills for evaluating the task and for using strategies optimally. Lower proficiency students likely have less control over the language and lack the appropriate background in the L2 to use some LLSs successfully (Vann and Abraham, 1990). "Knowing how to assess the success of a given strategy and apply corrective feedback to its use may be a more important skill to develop" (N. J. Anderson, 1991, p. 469).

7. Young L2 learners are able to use cognitively more demanding strategies as their L2 proficiency increases: Strategy use may follow a natural order, from LLSs that require no interaction (repetition, memorization, use of formulaic expressions) to LLSs that involve initiating and maintaining interaction (elaboration, request for clarification). Strategies that indicate awareness of and monitoring of errors seem to appear later in L2 development, though some learners' strategy repertoires do not expand and they often rely on the same early strategies later in the language learning process (Chesterfield & Chesterfield, 1985).

8. Proficiency level does not necessarily limit high school foreign language learners' use of LLSs: Exceptionally effective students at beginning, intermediate, and high proficiency levels use similar strategies, use them effectively, and use self-monitoring to determine whether their LLS use is successful. Highly effective beginning students are constrained primarily by limits in vocabulary and L2 knowledge, and not by an inability to use meta-cognitive strategies (Chamot and Kupper, 1989).

In summary, several useful observations come out of this research: (a) L2 proficiency does appear to be related to the types and frequency of strategies used by learners; (b) some learners may be better able to use meta-cognitive strategies as they develop greater control over the L2; (c) certain strategies appear more often among all learners and some strategies are more characteristic of learners at certain stages of L2 development; (d) success in the L2 may relate to how effectively and appropriately learners use LLSs; and (e) explicit LLS instruction may be necessary for learners who are less successful in their L2 development.

Proficiency and pronunciation strategy use. Much less has been done to investigate pronunciation strategy use, and even less in relation to L2 proficiency level.

Eckstein (2007) found that pronunciation scores were positively correlated with use of pronunciation strategies. Higher ability learners used pronunciation strategies more frequently than lower ability learners. However, limitations of these findings are similar to the general LLS research: The findings relied on correlational data and learner self-report.

Haslam (2010) explored the relationships among four factors: language aptitude, L2 proficiency, pronunciation strategy use, and learning context (EFL vs. ESL intensive-English programs). Overall language aptitude and learning context did not predict pronunciation gains over 10 weeks of instruction, but strategy use did. The author found a positive relationship between pronunciation strategy use and gains in ratings of comprehensibility and pronunciation accuracy (based on segmental accuracy only). No relationship was found between pronunciation strategy use and gains in global foreign accent or fluency. Haslam also observed that some learners with higher auditory aptitude scores used practice and noticing strategies longer (in terms of weeks) than did learners with low auditory aptitude scores. However, these findings were not statistically significant. She noted that high language aptitude learners used a greater

variety of strategies, but did not use any specific strategies more often than low aptitude learners. Again these results are based on correlational data and learner self-report.

With regard to proficiency, Sardegna (2009) found that, based on end-of-semester measures, low proficiency learners generally made smaller accuracy gains that higher learners. However, more than half of the learners with low entering proficiency were performing the same as the high proficiency group by time 4 (14 to about 42 months post-instruction). Those with high entering proficiency generally maintained their higher level of accuracy. This suggests that over time many (but not all) lower proficiency learners may catch up with higher learners. M. K. Hahn (2002), studying similar students in the same teaching context found that low proficiency students consistently performed lower than high learners at times 1 (pre-instruction), 2 (end of semester), and 3 (several months later).

Summary. The literature on general and pronunciation LLS use provides useful information about what we might expect of learners at different levels of L2 proficiency: Learners at all levels often rely on the same strategies, and, as proficiency develops, learners often use more meta-cognitive (self-monitoring, self-evaluation) and increasingly interactive strategies. Also, less successful learners tend to use a smaller group of strategies, use them less effectively, and often are less self-reflective. Thus we can look at learners as falling into, at minimum, two categories: (a) those for whom LLS use is effective, that is, advances in L2 proficiency seem related to effective and expanding use of LLSs, and (b) those who, in spite of active strategy use, are less successful L2 learners. Researchers of general LLSs use have found this pattern (Breun, 2001a, 2001b; Chamot & Kupper, 1989), as have those looking at pronunciation strategy use (M. K. Hahn, 2002; Sardegna, 2009). The current study does not

attempt to understand why such differences occur, but this issue is discussed again in Chapter 7 with regard how to this study's results offer implications for language teaching.

Because the current study involved a group of learners in an intact classroom, proficiency level could not be controlled. However, as is the case in most language classrooms, learners varied in noticeable ways in terms of proficiency. Thus, in this study, I explore differences in pronunciation strategy effectiveness according to entering pronunciation proficiency. Based on the previously mentioned findings, I expect that proficiency will be positively correlated with accuracy gains following strategy use. Additionally, I expect that for some learners, strategy use may not prove effective for increasing pronunciation accuracy and that learners may vary in terms of which strategies result in increased accuracy.

Summary and Future Research Needs on Pronunciation Self-Monitoring Strategies

Overall, the use of the self-monitoring strategies for learning L2 suprasegmental features has received limited attention in the LLS and pronunciation pedagogy literature. As noted earlier, many writers and theorists claim the importance of teaching L2 learners to self-monitor and self-correct their pronunciation. Strategy taxonomies for pronunciation learning and general language learning include self-monitoring, self-evaluation, self-correction, critical listening, and rehearsal as strategies used by learners (see Appendix C). A few studies have been completed using some of the targeted strategies and have reported improvement in learner pronunciation, but most of the studies contain design weaknesses that preclude claims that use of the activity resulted in positive pronunciation change. More systematic study is required to gain a better understanding of how the targeted strategies can be taught effectively to learners and the extent to which the observed use of these activities by learners results in increased L2 suprasegmental accuracy.

Chapter 3

Rationale for Focusing on Suprasegmental Features

In this chapter, I offer a rationale for selecting the phonological content for this study and present the following sections: the nature of suprasegmentals; a rationale for focus on suprasegmentals in the research literature; a rationale for suprasegmental focus in this study; and what is known about the effectiveness of suprasegmental instruction.

Description of Suprasegmental Features

In this part, I focus on descriptions of suprasegmental features, how they relate to segmentals, and the rationale for selecting the targeted phonological content for this study (message unit boundaries, primary phrase stress, intonation, reduction of unstressed syllables in content and function words, linking, word stress, construction stress). Additionally I discuss the targeted content's roles in L2 intelligibility, and instructional practices for improving learners' control of the targeted content.

The nature of suprasegmentals. When discussing the sound system of languages, researchers and linguists typically refer to two major elements: segmentals and suprasegmentals. As the following description indicates, the two are highly interrelated; however, for pedagogical and descriptive purposes, segmentals and suprasegmentals generally are treated as distinct categories. The segmentals, vowels and consonants, typically can be identified as discrete elements of speech that form words when they occur in meaningful strings. Consonants are typically described according to the extent to which air flow is constricted as it moves through the vocal tract (e.g., stops such as /d/ vs. continuants such as /z/), place of articulation (e.g., alveolar ridge) and articulators (e.g., tongue tip), and amount of vocal fold vibration (resulting in

homo-organically produced voiced vs. voiceless contrasts, such as /s/ and /z/ in English). Vowels do not involve vocal tract obstructions in the ways that consonants do, thus different descriptors are used: tongue height (low, mid, high), tongue fronting or backness (back, central, front), and lip position (spread, neutral, or rounded). For example, the English vowel /iy/, as in the word *feet*, is a high, front, unrounded vowel.

Suprasegmentals are those features that are not segmentals, including articulatory setting (Laver, 1980, cited by Clark & Yallop, 1995), which can be described as the typical vocal tract positions for a given individual or L1; pitch, which is the "perceived correlate of fundamental frequency" (p. 332); duration of segments or syllables; loudness, which is the "perceptual correlate of intensity" (p. 334); pitch patterns; and declination (the gradual decline in pitch from the beginning to the end of a span of speech). Prosody, which is the language-specific organization of suprasegmental features, has discrete features as well, though "prosodic phenomena *tend*, much more than consonants and vowels, to be directly related to higher levels of linguistic organization, such as the structuring of information" (p. 329).

Segmentals and suprasegmentals are not separate entities, rather they are integrated parts of a language's sound system. As an example, the segments of syllables receiving PPS or major word stress are longer in duration, typically higher in pitch, and often louder in relation to unstressed segments, and stressed vowels are full and not reduced. Vowel pitch is impacted by voicing of adjacent consonants. Adjacent sounds at word boundaries within a MU interact in ways that adjacent sounds separated by an MU boundary do not. Within a MU, linking, blending (palatalization), and elision of individual sounds may occur, whereas adjacent sounds separated by a MU boundary do not undergo these processes. One must know the intonation pattern of an utterance in order to predict likely effects at the segmental level (Clark and Yallop, 1995).

Clark and Yallop (1995) suggest a continuum of prosody, "ranging from the nonlinguistic or extralinguistic at one end [voice quality, characteristics of the speaker's vocal tract], through the paralinguistic [an in-between, gray area in which the speaker's intention or level of control over the features may be unclear to the observer], to the essentially linguistic [e.g., stress and tone]" (p. 329). A number of factors affect the prosodic structure of an utterance, including "focus, new vs. given information, beliefs about the assumptions shared by two conversing speakers, and quantitative factors, such as rhythm, number of elements and speaking rate", and morphosyntactic structure (Shattuck-Hufnagel & Turk, 1996, p. 233).

Next I briefly describe nine commonly designated suprasegmental features in English: MUs and their boundaries, primary phrase stress, intonation, rhythm, vowel reduction, linking, word stress, multiword constructions, and articulatory setting.

Message units. Also called breath groups, thought groups, intonation groups, or intonational phrases, among other terms, the MU (from L. D. Hahn & W. B. Dickerson, 1999) is a primary unit of spoken discourse consisting of a word or phrase, each having its own "meaningful tune" (Cruttenden, 1997, p. 7) or intonation pattern. In this study, the MU was chosen as the unit of analysis for two primary reasons. First, many phonological processes function at the level of the MU, resulting in specific phonetic outcomes, such as primary phrase stress, intonation patterns, and linking and palatalization at word boundaries. Second, L2 learners' creation of MUs is related to their comprehensibility. MUs that are too long are difficult for listeners to process and MUs that are too short may sound abrupt or too emphatic.

Message unit boundaries. MU boundaries usually are signaled by tones, pre-boundary lengthening of segments, and/or pauses (Shattuck-Hufnagel & Turk, 1996), though in spontaneous speech such phonetic cues may be absent (Cruttenden, 1997). The same utterance,

with different MU boundary placement (denoted by ｜), may have different meanings. For example,

> At Cornell, ｜ I studied math and physics for one semester.
>
> At Cornell, ｜I studied math ｜ and physics for one semester.
>
> (L. D. Hahn & W. B. Dickerson, 1999, p. 25)

In the first example, the speaker studied both subjects for one semester. In the second example, the speaker studied math for an unspecified time, perhaps during his or her entire time at Cornell. Though MU boundaries are determined by the speaker and typically relate to syntactic boundaries, their location is not predetermined or fixed.

Characteristics of targetlike English MU boundaries. In English, the end of a MU is signaled by completion of the intonation pattern that follows primary phrase stress. The final syllable in a MU may be lengthened, even if it is not stressed. A new MU may be signaled by an acceleration of the first few unstressed syllables of the MU. Pauses may or may not occur between MUs. No linking is used between the end of one MU and the beginning of the next. Targetlike MU boundaries typically occur following completion of grammatical phrases or complete ideas.

When determining target MU boundaries in this study, students' own patterns of MU length were respected. For example, some students tended to use longer MUs that were still within the target upper limit (see next paragraph), even though these longer MUs could have been divided into two or three shorter ones. As long as the MU boundaries fell within the criteria presented next, the MU boundaries were considered targetlike.

Following are criteria used in this study (and in its teaching context) for identifying appropriate MU boundaries (from W. B. Dickerson, personal communication, 2009):

1. Length: Students are instructed to produce message units that are approximately five to nine words in length, following research on short-term memory originated by Miller (1956), suggesting a limit of seven plus or minus two chunks of information. One- or two-word MUs are acceptable when each MU contains an element of a series of items or common short phrases, such as the phrase *for example.*

2. Keep together:
   Article + noun
   Adjective + noun
   Adverb + adjective/adverb
   Preposition + its object
   Auxiliary + verb head
   Verb head + particles
   Short subject + its predicate
   Short prepositional phrase + rest of phrase
   Same parts of speech joined by conjunctions

3. Break at the pause:
   Between long subject + its predicate
   Between prepositional phrases of four or fewer words followed by a longer string
   Between dependent and independent clauses
   Before conjunctions joining simple sentences

Production is considered nontargetlike when MU boundaries are placed in locations other than those listed above as targetlike. False starts, repetitions, self-repair, fillers, and hesitations that appear to be due to online processing or pauses to take a breath are items that do not indicate MU boundaries and thus are not counted as errors.

Primary phrase stress. According to the Clark and Yallop (1995) definition, PPS is included as a component of English intonation. In this study, PPS is treated separately, so that both the pitch pattern (fall, rise, fall-rise) and stress placement can be monitored separately by learners. In English, PPS includes lengthening of the vowel undergoing primary stress, a pitch move (jump or drop) on the stressed syllable, and is used to signal prominence of a word or syllable in a MU. Most often PPS occurs on the last content or function word in "new" information in a phrase ("old" or "given" information is not highlighted in spoken English), but it is also used to signal contrasts, contradictions, comparisons, choices, and other information the

speaker intends to highlight (L. D. Hahn, 1999, 2004). In the following example, old information appears within parentheses and PPS is signaled with a dot, •.

If the pH value is lower than seven, • | then it's an acid. • | (If the value is) larger (than seven), • | (then it's a) base. • (L. D. Hahn & W. B. Dickerson, 1999, p. 62)

In these next two examples, PPS is used to highlight contrasted information. The second example demonstrates how PPS can cause word stress to shift, as in *íncrease* and *décrease*, as well as the possibility of multiple PPSs in the same MU (the latter being an example of contrasts in parallel phrases, as described by L. D. Hahn & W. B. Dickerson).

This test has • construct validity | but not • face validity. (p. 20)

The temperature is going to • increase • today | and • decrease • tomorrow.

Characteristics of targetlike PPS. PPS is targetlike when it falls in an appropriate location, given the discourse structure. PPS is the most prominent syllable in a MU, signaled by pitch jump or drop (pitch change must be noticeable, in relation to surrounding syllables); syllables following PPS typically are quicker and quieter. PPS must be detectable and in the correct location to be considered targetlike. Errors include absence of PPS when it is expected, when it is used on the wrong syllable in the MU, or when multiple prominences are used and a single primary stress is not distinguishable. Multiple PPSs may appear in an MU that contains contrasted information.

Intonation. Within phonology, the definition of this term varies according to the theoretical position or the purpose of the writer. Ladd (1996) explains well the dual roles that intonation fulfills in a language such as English: first, a linguistic role, which implies that intonation has a phonological structure that can be described and explained. A second role is a

paralinguistic one, such that intonation acts like a "parallel linguistic channel" (p. 1) that can signal speaker traits such as age, gender, attitude, and emotion. Ladd's own phonological description indicates a separation of form and function. His goal is to identify the forms, that is, the phonological categories, and then to use those forms when describing how certain meanings are evidenced in speech; whereas other researchers and theorists start with the function (attitudes, meanings, emotions) and attempt to describe the forms, or intonation patterns, associated with a specific meaning.

Wennerstrom (1998) is interested in the use of intonation for cohesion in academic discourse. She provides a number of useful descriptors of intonation, drawing on key findings in the research literature. She first explains how intonation may function at both the phrase and word level. At the phrase level, phrase-ending intonation signals the relationship of the phrase to surrounding phrases: a rising intonation indicates nonfinality, that the listener is to interpret the clause in relation to what is to follow. Intonation also functions at a level above the phrase. An important discourse-organizing concept is that of the "paratone," which is much like the idea of the paragraph in written English. In a paratone, the pitch range is expanded at the beginning of a new topic and compressed at the end (Brown, 1977; Beckman and Pierrehumbert, 1986; and others, cited in Wennerstrom). According to Wennerstrom, pitch accent shows the status of lexical items in relation to the listener's knowledge, and it is used to show the relationship of the speaker's contribution in relation to the knowledge, ideas, and beliefs assumed to be shared with the listeners. For example:

- High pitch accent: used for new lexical items being added to the "mutual belief space" (Pierrehumbert & Hirschberg, 1990, cited in Wennerstrom) of the discourse.
- Low pitch or de-accent: used for items already believed to be part of the belief space.
- Contrasting pitch: a steep pitch peak used for contrasting items in the discourse.

Wennerstrom also describes how Halliday and Hasan's (1976) categories of cohesion can be related to pitch. For example, reference, substitution, and lexical cohesion (reiteration and collocation) are typically "given" or old information because they have antecedents in the text. Using Pierrehumbert and Hirschberg's concept of mutual belief space, the cohesive features described by Halliday and Hassan typically receive low accent because they are part of the knowledge shared by speaker and listener.

According to Brazil (1997), a goal in interaction is to increase the area of common ground between speaker and listener, to work toward social convergence. This is quite similar to Pierrehumbert and Hirschberg's "mutual belief space". Listeners generally expect that the speaker will provide new information that will increase common ground. So when speaking, the speaker alternates between known and new information in order to maintain social convergence. If information is presented as only new, then common ground decreases and social distance is increased.

In Brazil's model, however, the focus remains on the phrase and not the word level. The tone choice following tonic or primary phrase stress is used to indicate the common ground between speaker and listener. For example:

- Falling tones indicate new assertions.
- Rising tones denote shared knowledge.
- Level tones are used for information that is routinized; not necessarily shared or new information; or procedural information (referred to as oblique orientation).
- A level tone also may be used for pauses that occur as a person is processing what she wants to say (this is different from the function of the level tone in oblique orientation).

Speakers can use tone choice to highlight the common ground between the speaker and the hearer, thus building social convergence and decreasing the affective distance between speaker and hearer.

For the learners in the current study, understanding the various forms and functions of intonation is critical for accurate and clear oral communication in English. However, instruction condenses this theoretical background into three categories of intonation (fall-rise, fall, and rise, which are based on pedagogical aims and the text for the target audience: L. D. Hahn & W. B. Dickerson, 1999). Examples of each intonation pattern follow:

Instructor: I'm offering office hours on Monday, | Wednesday, | and Friday.

(fall-rise) (fall-rise) (fall)

Student: Did you say Monday?

(rise)

The use of fall-rise intonation indicates that the speaker likely will continue to speak, which is the case, since the instructor is listing the days for office hours. Falling intonation is used to signal that a thought is complete, and rising intonation is used for several question types in English, including repetition questions such as the one made by the student in this example. One can see how using incorrect intonation in these simple examples could lead to misunderstanding or increased processing time on the part of a NS listener.

Characteristics of targetlike intonation. In this study, intonation is considered targetlike when the appropriate pitch pattern is used following the PPS in a message unit: Use of rise to high range typically signals a particular question type; fall-rise to midrange signals nonfinality,

that is, the speaker intends to continue speaking; and fall to low range, signals completion of a thought. Absence of a distinguishable and appropriate intonation pattern is considered an error.

Rhythm. Rhythm in English refers to the characteristic alternations of stressed and unstressed syllables within phrases. Factors that contribute to the rhythm of a given language include syllable structure (number of segments in the onset and coda and nucleus length), contrastive vowel length, vowel reduction, vowel lengthening in specific contexts, and long vowels, such as diphthongs in English or nasal vowels in French (Clark & Yallop, 1995; Ramus, Nespor, & Mehler, 1999). The distinction often is made between two general language types: (a) languages such as English, Dutch, and Polish, which are said to have a "stress-timed" rhythm, in which stressed syllables are noticeably longer and more prominent and unstressed syllables are much shorter and quieter, and (b) languages which often are described as "syllable-timed," such as French and Yoruba (Cruttenden, 1997). The latter category of languages tend to have syllables of nearly equal composition and length, resulting in a more equal duration of each syllable.

Ramus, Nespor, and Mehler (1999) analyzed vowel/consonant temporal ratios and found that languages fall on a continuum of rhythmic variation, rather than within only a few rhythm categories (such as the syllable-timed and stress-timed dichotomy noted previously). Van Santen and Shih (2000) were interested in similar questions of the role of prosody in timing and found that in both English and Mandarin (languages that are typologically quite different), syllables exhibit "large and systematic variations in syllable duration" (p. 1020), even though the researchers controlled for syllable type and context. The segmental constituents of a syllable contribute more to syllable length than do any prosodic characteristics of the language (e.g., the rhythmic patterns typically described as stress-timed or syllable-timed). Yet for pedagogical purposes, the distinction between stress- and syllable-timed languages remains useful,

particularly in cases where the learners' L1 and L2 have very different syllable structures and vowel/consonant temporal ratios (e.g., English and Mandarin).

In English, content words such as verbs, nouns, adjectives, and adverbs receive heavy stress, as do "loud" function words (question words, demonstrative pronouns, and negatives). "Soft" function words (e.g., pronouns, articles, auxiliaries, prepositions, conjunctions) are typically unstressed (W. B. Dickerson, 1989). The result is a characteristic alternating rhythm which, for pedagogical purposes, often is described as isochronic: heavy stresses tend to occur at regular intervals (Clark & Yallop, 1995; Cruttenden, 1997). This isochronic pattern is most salient in English rhymes (○ for heavy stress, • for unstressed syllables):

○ • ○ ○ • ○

Hope of gain lessens pain. (Benjamin Franklin, *Poor Richard's Almanac*, cited in Hahn & W. B. Dickerson, 1999, p. 45)

In this example, the syllables *hope, gain, less,* and *pain* are spoken at regular time intervals and are given similar stress and length, whereas the two unstressed syllables are spoken quickly with shorter duration to maintain the isochronic pattern.

Two processes that occur in English to facilitate its characteristic rhythm are included in the teaching context investigated in this study: vowel reduction (in unstressed syllables or words) and linking (Clark & Yallop, 1995; Cruttenden, 1997). Following are brief descriptions and examples of each process.

Vowel reduction. Stressed and unstressed vowels in English words differ in four key ways. Stressed vowels are full vowels, longer in duration, usually higher in pitch, and usually spoken with greater intensity. Unstressed vowels contrast with stressed vowels in that they are reduced to /ə/, as in the second syllable of *custom*, or /ɪ/, as in the second syllable of *visit*, though quick full vowels are used in the final unstressed syllables of *radio*, *menu*, *happy*. Unstressed

syllables are spoken very quickly and more quietly, and often at a lower pitch, than nearby stressed syllables. This contrast between stressed and unstressed syllables is critical for creating the characteristic rhythm of English. In a phrase such as *When will he arrive*? *When* and the syllable *rive* are the only stressed syllables in this phrase; the others are reduced.

Characteristics of targetlike vowel reduction. Vowels in unstressed syllables are quick and reduced, and vowels in stressed syllables are full and longer in duration. Using a full vowel instead of the correct reduced version is considered an error. For example, a student may produce the word "method" using a full vowel in the both syllables, instead of reducing the unstressed vowel in the second syllable to /ə/.

Linking. Within MUs in American English, pauses do not occur between words, rather all words are connected to each other. Also, when the end of a word is in contact with the beginning of the next, other articulatory processes occur, such as resyllabification, co-articulation, and assimilation. In the following sentence, underlining is used to show how the consonant at the end of one word links to the vowel at the beginning of the next: *A numbe<u>r o</u>f processe<u>s o</u>ccu<u>r i</u>n <u>E</u>nglish.* Linking also occurs when adjacent words in a MU end and begin with the same sound, as in *produ<u>ce s</u>ounds*. No pause occurs between the two words and the interword /s/ may be lengthened. When adjacent words end and begin with vowels, as in *m<u>y o</u>nly*, the off-glide of the vowel in *my* (in this case, /y/) links to the onset vowel of *only*. When the adjacent sounds (in this case, within a word) are stops, as in *pop<u>c</u>orn*, the final /p/ is unreleased before the /k/ is articulated. Lack of linking within MUs disrupts rhythm by breaking the smooth connections between words in a MU. Sometimes such breaks are used intentionally in English to denote emphasis, as in "Did you say / yes / or no? But if linking is missing frequently in the speech stream, the result is a nontargetlike rhythm that decreases the speaker's intelligibility and

may also negatively influence a NS listener's attitude toward the NNS (non-native speaker) (Anderson-Hsieh et al., 1992; L. D. Hahn, 1999, 2004).

Characteristics of targetlike linking. Within MUs, word-final sounds are linked to the beginning of the following word, either as consonant-to-vowel, vowel-to-vowel, or same/similar consonant to same/similar consonant. The following are considered errors: A pause between words within a MU, epenthesis (syllable- or word-final vowel-insertion), or a glottal stop preceding word-initial vowels or semi-vowels.

Word stress. In English, one syllable within a multisyllable word is more prominent than the other syllables. This prominence is signaled primarily by a pitch change, but may also be accompanied by increased loudness and duration relative to surrounding unstressed syllables. Vowel quality also plays a role in signaling prominence, with unstressed syllables typically containing reduced vowels and stressed syllables containing full vowels (Clark & Yallop, 1995). Using these descriptions of stress, a listener could expect when hearing a two-syllable word such as *model* that the vowel of the first (stressed) syllable will be longer in duration and likely a bit louder than the vowel of the second (unstressed) syllable; that perceptibly different pitch levels will be used for the first and second syllables; and the first vowel will be a full vowel and the second reduced to /ə/ or a syllabic /l/. However, L2 speakers of English commonly stress both syllables equally, producing the full vowel /ɛ/ in the final syllable of *model*. The lack of contrast in syllable length and vowel quality between the stressed and unstressed vowels may significantly decrease the intelligibility of this word for a NS listener.

Though L2 learners (and many language instructors) often perceive English as having no pattern for assigning word stress, rules do indeed exist that can be used for predicting stress location in the great majority of polysyllabic content words in English. In order to use the rules,

learners must have knowledge of syllable structure and be able to identify parts of speech and suffixes and prefixes (see W. B. Dickerson, 1989, and L. D. Hahn & W. B. Dickerson, 1999, for complete explanations of the word-stress rules taught in the course studied in this dissertation).

Characteristics of targetlike word stress. The key features for determining targetlike word stress for this study include the following: The correct syllable receives the major stress, and the stress is perceived as a distinguishable prominence. Errors include no distinguishable stress (all syllables in a word are perceived as equal in prominence); multiple prominences of equal value; or stress on the wrong syllable. Use of a full vowel in an unstressed syllable does not necessarily indicate an error, as long as the syllable that should be stressed is clearly more prominent due to vowel length and/or pitch or intensity contrasts.

Multiword construction stress. In English, groups of two or more words often function as a single part of speech. These constructions have a special meaning when the words appear together. Some of these are referred to as "compounds." Such constructions can be compound nouns such as *blackbòard*, compound verbs such as *déadlòck*, compound adjectives such as *práise-wòrthy*, and compound numbers such as *four hundred and ninety-síx*. Multiword constructions have one prominent stress, often (though not always; compound numbers follow a different pattern) on the stressed syllable of the penultimate word in the construction. As with word stress, stress on multiword constructions often is predictable, though the system is complex (Bolinger, 1986; L. D. Hahn, 1994). Stress placement may affect the meaning of the same two-word construction. For example the *Whíte House* is a specific building, whereas a *white hóuse* (an adjective + noun construction) refers to a house that happens to be white in color. Thus multiword stress is an important pronunciation feature for L2 learners to understand and produce accurately.

Characteristics of targetlike multiword construction stress. Features of targetlike construction stress are similar to those for word stress and include the following: Detectable stress is present on the correct syllable in multiword constructions. The stressed syllable contains a full vowel. A pitch jump or drop may be used on the stressed syllable (pitch change must be noticeable, in relation to surrounding syllables). Errors include absence of or incorrect placement of stress.

Articulatory setting. Articulatory setting refers to the overall vocal tract conditions that influence production of segmental and suprasegmental language features. These conditions can be unique to the speaker or may represent overall traits for a given language, such as greater lip-rounding in French and open jaw and spread lips in English (Esling & Wong, 1983). Using the articulatory settings from one's L1 when speaking an L2 is one factor that contributes to perception of a "foreign accent." Thus a challenge for L2 learners who want to minimize the perceived strength of their accent is to make articulatory adjustments that correspond to L2 settings.

This suprasegmental feature remains poorly understood from research and pedagogical perspectives and thus is not a focus of the current study. Monitoring articulatory setting is difficult because one cannot directly observe most articulatory features without disturbing an articulation. A goal of this study was to focus on suprasegmental cues that learners can most easily identify and that require readily available resources for monitoring. Articulatory setting does not fulfill these criteria and was not included in this study.

Summary. In this section, I introduced the suprasegmental features commonly discussed in the pronunciation pedagogy and research literature. I described how segmentals and suprasegmentals are interrelated and briefly defined message units, message unit boundaries,

primary phrase stress, intonation, rhythm, vowel reduction, linking, word stress, multiword construction stress, and articulatory setting. In the next section I discuss why suprasegmentals are worthy of study and provide a rationale for selecting the suprasegmentals included in the proposed study.

Rationale for Focus on Suprasegmentals in the Research Literature

Suprasegmentals are worthy of study for several reasons. First, most adult ESL learners have trouble producing these features accurately and typically apply their L1 prosody to their spoken English (Swan & Smith, 2001). Second, these features can be operationalized in a way that makes them salient for learners: They can listen for specific features (pitch move, loudness, duration, vowel quality) in their own and in others' speech. Third, these suprasegmental features contribute heavily to intelligibility in English, and the ESL learners targeted in this study must strive for high intelligibility if they wish to be successful as teachers in an English-language university classroom. I describe these three rationales in more detail in the following sections.

Difficulties learners have in acquiring English prosody. The literature on pronunciation instruction and research commonly cites the difficulties adult learners have in acquiring L2 prosodic features. Research on ESL speakers from L1s that rely heavily on tonal cues rather than on English-style stress cues often are more successful producing English pitch and intensity cues, which are relevant cues in the L1, and less successful producing English duration and vowel quality cues, which generally are less relevant L1 cues (Hua, 2003; Nguyen & Ingram, 2005).

An implication of these phonological differences among languages is that NSs of English, who are reliant on stress cues for word recognition, have difficulty processing L2

speakers' production of English when stress cues such as contrasts in duration, pitch, intensity, and vowel quality are not produced in a manner consistent with native English production. Research on human speech perception has indeed demonstrated this. Benrabah (1997), citing the author's and other researchers' data, found that when native English listeners transcribed the speech of Arabic speakers of English, they often misperceived words that were incorrectly stressed. For example, the word *normálly* (stressed incorrectly on the second syllable) was perceived as *no móney*. These research findings indicate that listeners were more reliant on stress cues than on segments (or context) for speech recognition.

Second-language learners typically have difficulty learning to produce English word stress (Anani, 1989; Archibald, 1995; Avery & Ehrlich, 1987; Aziz, 1980; Baptista, 1989; Cooper, Cutler, & Wales, 2002; W. B. Dickerson, 1978; Field, 2005; Swan & Smith, 2001). The few studies that have looked at word stress production by L2 speakers of English have shown interference from L1s that have different stress cues or rules or that do not use lexical stress (Benrabah, 1997; Guion, Harada, & Clark, 2004; Kawagoe, 2003). Mandarin Chinese speakers also are reported to have difficulties with English word stress. Y. Chen (2001) investigated the acoustic features of English word stress produced by Mandarin speakers and found that their use of pitch, duration, and intensity were similar to that of NSs. However, they produced greater extremes of pitch and used higher pitch and greater vowel duration on unstressed syllables than NSs. Juffs (1990) noted that Chinese speakers of English evidenced frequent stress errors, in terms of placement and the phonetic processes used to signal it (though he does not provide details about these errors). Due to the importance of word stress in comprehensibility, L2 learners of English would benefit from understanding how to produce English word stress cues correctly.

Swan and Smith (2001) summarize for language teachers the most common pronunciation problems experienced by English learners from eight language groups. They note that all language learners will require instruction in English stress and rhythm due to the seemingly complex nature of English stress. Common rhythm problems include difficulty reducing unstressed vowels; constraints on syllable types of an L1 such as Japanese (e.g., only open, or consonant-vowel [CV], syllables are allowed) may also cause learners to insert vowels at the end of closed (CVC) syllables, thus changing the stress pattern of the word and phrase; and learners from syllable-timed languages may stress all syllables equally, unaware of the importance of the duration cue for stressed syllables.

Some research has suggested that the L1 can facilitate production of the relevant English word-stress features. Nguyen and Ingram (2005) found that Vietnamese speakers were able to effectively use pitch as a cue in English, drawing on their knowledge of pitch use in the L1. Cooper, Cutler, and Wales (2002) found that Dutch speakers were able to use their L1 sensitivity to word stress cues to identify stressed English words based on hearing only the first syllable. Cutler, Weber, and Otake (2006) and MacKay, Flege, Piske, and Schirru (2001) discuss how L2 speech sounds are often equated with similar L1 speech sounds, which can block the formation of the L2 phonetic category, though such blocking does not prevent phonetic learning from occurring. The tone use in languages such as Japanese, Chinese, and Vietnamese and pitch use in English for signaling word stress may be an example of such equivalence classification. Though this does not suggest that L2 speakers cannot learn English word stress cues, the concept of equivalence classification may explain why L2 speakers have difficulty achieving native-like production of these cues.

Several researchers note the problems L2 speakers have with English intonation and rhythm (Perez-Gamboa, 1989; Pickering, 2001; Ramirez Verdugo, 2006; Wennerstrom, 1998). Because of the difference in syllable structure and the different function of tone in both languages, Chinese ESL speakers tend not to link words in MUs and may try to apply tones to individual syllables or words, resulting in a nontargetlike prosody (J. Chang, 2001).

Because adult ESL learners typically encounter the types of difficulties with English prosody that have just been described, the targeted content is a primary focus in this study's teaching context.

Salience of targeted content. As noted in an earlier section on teaching perception, prediction, and production of the targeted content, before learners can monitor their pronunciation more holistically, they are trained to listen for the distinct cues that are considered most important for signaling each target feature. Rather than providing the general instructions to "listen for" or to "produce" PPS, they are trained to listen for or produce the phonetic cues of pitch change, syllable duration, and loudness. Additionally, the targeted content is described simply, using language that is accessible to students who generally do not have a linguistics background and who may never have thought about the discrete features of L2 phonology.

Learners often have a very basic familiarity with concepts such as pitch, stress, melody, and phrases, though these must be defined to ensure a common understanding. As previously explained, stress and pitch are evidenced differently from one language to the next and pitch may be used for different purposes. The cues used to perceive stress or pitch in one language may not be the same in another. Thus L2 learners of English likely do not know, without explicit instruction, which cues to attend to when speaking and listening in English. They must learn to distinguish pitch jumps from pitch drops, rising or falling intonation, and differences in syllable

duration. Content is divided into a few categories, making it easier to remember and apply during practice. For PPS, the pitch jumps or drops, and the most common predictive rules are presented in easy-to-understand formats. Intonation follows only three patterns: the pitch rises, falls, or falls and then rises. MU boundaries are perhaps the easiest for learners to grasp due to the salience of pauses, restarts, self-repairs, and other disfluencies present in their L2 production.

The role of targeted content in promoting L2 intelligibility. The pedagogical and research literature has a fair amount to say about the role of suprasegmentals in promoting L2 intelligibility. Though a consensus does not exist regarding which suprasegmental features have the greatest impact on intelligibility, research has indicated that each target feature contributes in its own way. In the following paragraphs I first describe how intelligibility has been defined and then offer a brief review of the available literature on the role of the targeted suprasegmentals in L2 intelligibility for ESL speakers.

Intelligibility and comprehensibility. Researchers looking at L2 speakers' pronunciation errors frequently refer to the impact of these errors on the speaker's intelligibility and comprehensibility. Fudge (1984) notes that "comprehensibility depends on rhythm, and therefore the placing of stress within words can play a large part in determining how well a native English hearer will understand the foreign speaker" (p. 4). W. B. Dickerson (1989) defines intelligibility as "speaking so that one is understood" and defines comprehension as "understanding when one is spoken to" (p. xii). An L2 speaker of English must produce accurate rhythm and stress to ensure that he or she is understood by the listener. Similarly, L2 speakers must become competent listeners or perceivers of English stress so they may comprehend NSs. Dickerson emphasizes that "to be understood, the speaker must use patterns of stress and vowel quality that native listeners are accustomed to hearing" (p. xii).

According to L. D. Hahn's (1999) review of the literature on intelligibility, researchers are not consistent in their use and definition of the term. She notes that some researchers define intelligibility much like W. B. Dickerson (1989): as "simple recognition of a word or utterance" (p. 54, based on work by Smith & Nelson, 1985). If a listener can transcribe or repeat accurately what the speaker said, then the utterance is intelligible. Similar to Dickerson, Smith and Nelson also distinguish comprehensibility from intelligibility and define the former as understanding the meaning of an utterance. Hahn notes that researchers frequently use the two terms interchangeably, as she does in her own research. I found similar interchangeability of the terms in the research literature. For the purposes of this study, I, too, will use the terms interchangeably, in that learners typically are concerned with understanding the speech of NSs and in being understood by NSs. Additionally, suprasegmental errors likely contribute differently to intelligibility and comprehensibility, with stress and rhythm errors likely having the greatest impact on the former (understanding what the speaker *said*) and errors in PPS, MU boundaries, and intonation impacting the latter (understanding what the speaker *meant*).

Speaker errors have an impact on the listener in other ways. L. D. Hahn (1999) cites Ludwig's (1982) observations that speaker errors can cause listeners to become focused on the form of the utterance rather than on the meaning of the message itself. Fayer and Krasinski (1987; cited in L. D. Hahn) further note that listeners are distracted and sometimes annoyed by speaker errors. Hahn points out the importance of the listener's role in determining intelligibility, that this is not an absolute characteristic of the speaker.

Bent and Bradlow (2003) specifically studied this issue (using samples of filtered speech) and found that definitions of intelligibility varied according to the interlocutors: NNSs of English from the same L1 background understood each other better than they understood NSs of English.

NNSs of English from different L1 backgrounds understood each other as well as they understood NSs of English. And predictably, NSs of English did not understand NNSs as well as they understood other NSs. They concluded that "any measure of speech intelligibility must take into account both talker- and listener-related factors" (p. 1608). Munro, Derwing, and Morton (2006), using unaltered speech samples found that L1 did not have as significant an impact on intelligibility, and that listeners from a variety of L1s generally agreed on how intelligible NNSs were, suggesting that the interlocutors' L1s were perhaps less important and the speaker's overall speech properties played the key role (i.e., their pronunciation). Jenkins (2002) also emphasizes how the mix of interlocutors greatly influences which pronunciation features are most important for intelligibility when speaking English as a *lingua franca*.

Thus for international teaching assistants (ITAs) teaching in an ESL setting, not only does their intelligibility relate to how targetlike their production is of PPS, intonation, word stress, and MUs, but also to factors within the context, such as how familiar the listener is with the speaker's voice or accent, level of familiarity with the lecture content, how willing the listener is to try to comprehend the ITA, and the listener's subjective reactions to a NNS's accent (Levis, 2005; Munro & Derwing, 1995).

Intelligibility and suprasegmentals. A number of studies point out the potential relationship between intelligibility and accurate use of prosodic features in English.

Tajima, Port, and Dalby (1997) demonstrated the negative impact of non-native prosody on NSs' word recognition. The researchers temporally corrected speech samples from Chinese speakers of English to match NS of English timing and corrected NS of English speech to match Chinese timing. They found that

> intelligibility of the unmodified Chinese-accented phrases was poor (39% correct), but improved significantly (to 58%) after temporal correction. Performance on the

> native productions was high (94%), but declined significantly (to 83%) after temporal distortion according to the Chinese speaker's timing. (p. 1)

In one example, the intended phrase *equal size* was identified as *you're concise* (p. 6)

Tyler, Jeffries, and Davies (1988, cited in L. D. Hahn, 1999) found that ITAs who used too many pauses, too many PPSs within the same MU, and falling intonation at inappropriate times were rated by undergraduates as "disorganized and unfocused" (p. 55). Constantinou (1993, cited in L. D. Hahn, 1999), in a study of the English speech of Mandarin speakers, found evidence for a high correlation between native-like prosody (based on acoustic measurements of duration and peak amplitude) and intelligibility ratings by NSs. Other suprasegmental features that Hahn identified in her literature review as contributing negatively to intelligibility were interstress intervals that were too long in comparison to NSs (P. Anderson, 1993), incorrect word stress, and incorrect PPS (Bansal, 1969, cited in L. D. Hahn).

Researchers have studied the impact of segmental and suprasegmental accuracy on the intelligibility of L2 speakers of English and have generally found that prosodic features, including stress, rhythm, intonation, and phrasing, are of great importance for successful communication, likely contributing more than segmental accuracy (Anderson-Hsieh & Koehler, 1988; Anderson-Hsieh et al., 1992; L. D. Hahn, 1999, 2004; Johansson, 1978; Pickering, 2001; Wennerstrom, 1998). Jilka (2000) studied NSs of German and English and perceptions of foreign accent (which is different and not necessarily directly related to intelligibility: Munro & Derwing, 1995) and found that intonation was more important than rhythm and speaking rate, but that segmental foreign accent was more important than intonation in accent perception (intonation was defined as the F0 contour). When investigating intelligibility as a measure of pronunciation proficiency among ITAs at a Canadian university, Isaacs (2006) found that

nontargetlike sounds and word stress contributed the most to negative intelligibility ratings and sentence rhythm and pitch played secondary roles.

Anderson-Hsieh et al. (1992) investigated how deviance in the use of prosody, pronunciation of segmentals, and syllable structure affect NSs' impressions of NNSs' speech samples. Basing their conclusions on SPEAK text readings from 60 male NNSs of English from 16 different L1 backgrounds, they found that accuracy of prosody (word and phrase stress, rhythm, intonation, phrasing, overall prosody) correlated significantly ($r = .90$, $p < .0001$) with overall pronunciation scores. They also found that prosody was more strongly correlated with pronunciation scores than were segmental or syllable errors. One concern with this study is that the ratings were based on readings rather than on naturally occurring speech. The intonation of text reading typically differs from that of spontaneous speech (L. J. Dickerson, 1975; Levis & Pickering, 2004; Tarone, 1985).

As mentioned previously, numerous studies have demonstrated the importance of lexical stress in auditory word recognition in English (Benrabah, 1997; Cutler, Dahan, and van Donselaar, 1997; Field, 2005; Slowiaczek, 1990). Results from research by Slowiaczek suggest that stress cues (duration, pitch, intensity, vowel quality) as well as segmental information contribute to the processing of words by NSs of English, and that L1 English speakers access lexical representations faster when input is correctly stressed.

A study by Field (2005) found that NS listeners of English had difficulty identifying incorrectly stressed English words, and the greatest impact on intelligibility occurred when stress was incorrectly shifted to the right (60% correct) rather than to the left (79% correct). Cutler and Carter (1987; cited in Field, 2005) note that 85.6% of English content words in running speech are monosyllabic or stressed on the first syllable. So it seems logical that a stress shift to the right

would change the perception of a word boundary, whereas incorrect stress on the left syllable would have a lesser effect because it still is signaling the start of a new word.

The type of intonation use by graduate students preparing to become ITAs has received attention recently. Wennerstrom (1998), in a study of 18 Mandarin speakers, found that accurate use of intonation, particularly the paratone, had a positive, significant relationship with SPEAK test scores of comprehensibility. Fagundes' (1994) study of 400 international graduate students found that scores on pronunciation (phonemic errors, foreign stress and intonation) and fluency (appropriateness of pauses) correlated highly with SPEAK test scores (pronunciation: $r = .92$, $p < .01$; fluency: $r = .89$, $p < .01$). Pickering (2001), in a study of six ITAs (L1 = Mandarin) and an analysis of their teaching presentations, found that, in comparison to NS TAs, the ITAs made less use of rising tones (used to signal convergence, or common ground in the discourse) and overused falling and level tones, which gave a monotonous feel to their discourse and created a sense of distance with their listeners.

PPS has been found to impact comprehensibility and attitudes toward NNSs. L. D. Hahn (1999, 2004) investigated undergraduate NS reactions to NNS monologues in which PPS was used correctly, incorrectly, or was absent. She found that "the participants recalled significantly more content and evaluated the speaker significantly more favorably" (p. 201) when PPS was used correctly than when used incorrectly or if it were absent.

Rationale for Selection of Suprasegmental Features for This Study

Research has offered evidence for the greater impact of suprasegmental accuracy as compared to segmental accuracy on NNS intelligibility and comprehensibility and on the reactions of NSs listening to NNSs (Anderson-Hsieh & Koehler, 1988; Anderson-Hsieh et al.,

1992; L. D. Hahn, 1999, 2004; Johansson, 1978; Pickering, 2001; and Wennerstrom, 1998). Students in the current study also tended to have mastered many of the most important English segments (based on functional load as defined by Brown, 1988) and most often needed help with message units, rhythm, primary phrase stress, multiword constructions, word stress, and intonation. All eight suprasegmental features were taught in this study, as were the segments that have a high error rate and high functional load in student diagnostics. Analysis of learner errors indicated that suprasegmental errors were common and thus their improvement should be most critical for future teaching effectiveness (e.g., by properly focusing their students' attention through correct use of PPS, by signaling social convergence through appropriate use of intonation, by facilitating listener processing through use of appropriate message unit breaks and accurate word stress).

To summarize, recent research has offered support for the importance of accurate use of suprasegmentals in promoting L2 intelligibility. The use of too many pauses and inaccurately placed MU boundaries (Tyler et al., 1988, cited in L. D. Hahn, 1999); absent or incorrect PPS (L. D. Hahn, 1999, 2004); nontargetlike intonation (Pickering, 2001; Wennerstrom, 1998); and word stress errors (Benrabah, 1997; Guion et al., 2004; Kawagoe, 2003) all have been shown to negatively impact L2 intelligibility. All are cues used by NSs of English, both for sentence- and word-level processing and for interpreting the meaning of utterances within discourse. ITAs commonly work with undergraduate populations who are unfamiliar with accented English, thus achieving a sufficient level of L2 intelligibility is critical for ITA academic and professional success. I have not found research that has definitively identified one of these features as most important. However, a definitive answer may not exist. Instead, what is important for

intelligibility most likely depends on a combination of speaker and listener characteristics, the speaking context and its communicative demands, and the frequency of the L2 speaker's errors.

Effectiveness of Suprasegmental Instruction

In the following section, I briefly review 22 studies that have investigated various types of suprasegmental instruction in laboratory and classroom settings (see Table 4 for summary). Four studies found that instruction was not effective (Ewing, 2002; Gorsuch, 2001; Harris, 2003; Macdonald, Yule, & Powers, 1994). Of the remaining 18, only eight included appropriate levels of control, thus allowing stronger claims about instructional effects. Many instructional practices are represented within this body of research and most studies used a mix of teaching practices, in which the purpose was not to isolate effects of a particular practice. Instructional practices included mixed practices (Anderson-Hsieh, 1990; M. G. Chang, 2006; Couper, 2003, 2006; Gorsuch, 2001; M. K. Hahn, 2002; Kendrick, 1997; Ramirez Verdugo, 2006), focus on suprasegmentals only (Akita, 2005; Aufderhaar, 2004; Couper, 2006; Derwing & Rossiter, 2003; Moyer, 1999; Pennington & Ellis, 2000) or segmentals only (Akita, 2005; Derwing & Rossiter, 2003), communicative practice (Macdonald et al., 1994), tracking and imitation of NS models (Harris, 2003; Myers, 1995), drills (Macdonald et al., 1994), use of speech visualization software (Anderson-Hsieh, 1992, 1994; deBot, 1983; Seferolug, 2005), language labs (Macdonald et al., 1994), oral reading (Ewing, 2002), focus on word stress only (Murphy, 2004), and other traditional classroom approaches that may have been given brief mention.

The study descriptions are organized into the following categories: awareness-raising for pronunciation features, overall pronunciation ability, PPS, word stress, intonation, and rhythm.

Within each category, studies are presented from oldest to newest. I did not find research focused only on instruction for improving MUs and multiword constructions.

Table 4

Summary of Research on Effectiveness of Instruction on L2 Suprasegmental Features

Suprasegmental instruction is NOT effective		
Study	Controls?	Key findings
Macdonald et al. (1994) (L) 10 minutes; 2-day post-test	Yes	No difference was found in effectiveness of four different short-term instructional methods.
Gorsuch (2001) (C) 38 hours	No	A textbook test showed no pronunciation improvement following instruction using a variety of teaching methods; the test used may not have content validity.
Ewing (2002) (C) 5 weeks, 25 minutes per day	No	Individual training (modeling, imitation, oral reading) on use of PPS and intonation is not effective.
Harris (2003) (C) 8 weeks, 1 hour per week	Yes	
Suprasegmental instruction IS effective		
Study	Controls?	Key findings
Aufderhaar (2004) (C) 8 weeks	No	Listening to and analyzing NS prosody results in stress/unstress ratios becoming more native-like.
Pennington & Ellis (2000) (L) short duration	No	Instruction (explanation, listening, oral practice) results in improved perception of L2 suprasegmental features.
M. G. Chang (2006) (C) 8 weeks + 2-month post-test	Yes	
Anderson-Hsieh (1990) (C) 10 hours out of 45 total	No	Use of a broad range of pronunciation instruction practices results in improved production accuracy.
Kendrick (1997) (C) 9 months	Yes	
Moyer (1999) (correlational study)	No	
Couper (2003) (C) 16 weeks	No	

(continued)

Table 4 (continued)

Suprasegmental instruction IS effective		
Study	Controls?	Key findings
Couper (2006) (C) 12 weeks	No	
Ramirez Verdugo (2006) (L/C) 10 weeks	Yes	
Myers (1995) (C) 16 weeks	No	Tracking results in higher speech ratings.
Akita (2005) (C) 4 months (90 min x 12 sessions)	Yes	Both segmental and prosody groups improved in perception/production of prosody; prosody group improved more
Derwing & Rossiter (2003) (C) 12 weeks, 20 minutes per day	Yes	Global instruction (focus on prosody) has more impact on accuracy and fluency than does segmental instruction.
Anderson-Hsieh (1992, 1994) (L) 6 to 8 hours	No	Use of speech visualization using VisiPitch or similar software for instruction for prosodic features is effective.
deBot (1983) (L) 60-90 minutes	Yes	
Seferolug (2005) (C) 3 weeks	Yes	
Murphy (2004) (C) one semester	No	Results are based on learners' self-reports that instruction on word stress was helpful.
M.K. Hahn (2002) (C) 16 weeks; post-tests 3 to 13 semesters later	Yes	Learned primary phrase stress patterns persist 3 to 13 semesters post-instruction.

Note. L = laboratory; C = classroom or individual instruction; amount of instruction provided

Awareness raising for pronunciation features. M. G. Chang's (2006) study followed eight Mandarin speakers during an 8-week ESL pronunciation class focused on intonation and

stress. She looked at how pronunciation was taught (how much class time on each target, type of T-S interaction, types of activities); how much students noticed the pronunciation features being taught, based on their comments from interviews, and she looked at T1 (pre-instruction), T2 (immediately post-instruction), T3 (delayed post-test 2 months later) speech samples to see if students implemented what was taught. Most pronunciation instruction was in a presentation-practice format and focused on intonation and also to a lesser extent on other suprasegmentals. Class time was evenly split among perception, production, and combined activities. She transcribed phonetically and compared student speech to what she would have done. Her rating of intonation was holistic (was it native-like) rather than evaluating it based on pitch move or direction. Chang found that learners' perception of the suprasegmental features improved but there was little evidence of change in production.

Effectiveness of instruction for overall pronunciation ability. Anderson-Hsieh (1990) gave ITAs 10 hours of suprasegmental instruction (out of a 45-hour course), using a variety of instructional practices, including classroom instruction on stress, rhythm, and intonation; perception practice; visual monitoring of suprasegmentals using VisiPitch; analysis of texts for MUs, PPS, word stress, and intonation; and voice recordings. The author offered many examples of the materials used. She found that SPEAK test scores increased by 33 points, though she cannot attribute this gain to the instruction due to the lack of a control group. No information on significance of the increase was offered.

Anderson-Hsieh (1992, 1994) describes how to use a specific product, VisiPitch, to provide electronic visual feedback when teaching English rhythm, linking, stress, and intonation, and reproduced in her article visual evidence of how ITAs improved their production after only 6 to 8 hours of instruction. She also provided extensive advice on how to accurately interpret the

technology and discussed how the L1 of the learner and the target language influence how the readings are interpreted. The purpose of this paper is not to evaluate the teaching method and she does not provide experimental evidence for its effectiveness.

Macdonald, Yule, and Powers (1994) set out to test the effectiveness of four instructional interventions commonly used in language classrooms and reflective of differing teaching methods and theoretical frameworks, but found no significant differences among the interventions, and for nearly all participants, accuracy deteriorated or remained unchanged. The focus was on pronunciation of targeted vocabulary in spontaneous speech. The participants were 23 adult Chinese L1 speakers (Mandarin), who provided an extemporaneous speech sample (T1) and then received one of the four interventions: (a) traditional in-class drills (participants met for one 10-minute session with a teacher, who provided feedback); (b) self-study with tapes (participants completed one 30-minute session alone in a language lab); (c) interactive activities (participants completed one 10-minute session with an instructor, who asked different clarification requests, but gave no feedback); and (d) no intervention (control), but participants were given 10 minutes for silent review. Immediately after the intervention, a second speech sample was elicited (T2). Two days later (T3), participants provided a third speech sample.

A primary limitation of this study is that the interventions and time elapsed were minimal: 10 minutes for three of the groups and 30 minutes for the self-study group; persistence of learning was measured two days later. The drill group received teacher feedback, but the others did not. Other studies provide training over a longer time period, and in many classroom settings, it would be expected that improvement would not be immediate, but take weeks, months, or longer. The authors do provide a useful rationale for why pronunciation deteriorates: that temporary decreased performance may indicate that restructuring is taking place.

Both W. B. Dickerson (1978) and Morley (1991) comment extensively about the amount of time needed for learners to improve and how teachers should interpret changes or lack of changes in student pronunciation. Certainly Macdonald et al. (1994) echoed this in their conclusion when they stated "that the individual learner may represent a more powerful variable than does the instructional setting in the acquisition of pronunciation" (pp. 95-96).

Myers (1995) studied the effects of a particular type of pronunciation instruction (tracking) on the SPEAK test scores of 11 international graduate students taking a pronunciation class in preparation to teach at a U.S. university. Two thirds of the students spoke Chinese as their L1; the others were from Korea, Turkey, Pakistan, or India. The class met 2 hours per week for 16 weeks. During each class students listened to recordings of textbook readings and mimicked the reader's pronunciation. Students reported that they felt the instructional method helped their pronunciation of words and phrases in a context, and SPEAK test scores improved by 43 points ($p < 0.01$) from the beginning to the end of the semester. Though this study did not include a control group (which would have offered stronger evidence of the effects of the particular type of instruction), it demonstrates the effectiveness of a classroom-based teaching approach and uses a reliable measure of speaking proficiency, the SPEAK test. In this study, students would have been mimicking accurate word stress, rhythm, and intonation, all of which contribute to SPEAK test scores.

Kendrick's (1997) study was motivated by a perceived lack of definitive evidence that pronunciation instruction results in long-term learning. Her study followed the progress of eight teenage NNSs of English enrolled in an English boarding school (an immersion setting). The students received a broad range of instruction over 9 months, some focused on specific phonemes using minimal pairs and focusing on articulation, use of rhymes and other techniques

for focusing on rhythm, use of drama, role play, mimicry and drills focusing on word stress and phrase rhythm, and use of individual study with tapes.

After six tests of progress, all students showed improvement and Kendrick found high correlation between instruction time and scores. Greater improvement was noted for phonemes than prosody. Scores on mimicry tests seemed to be the best predictor of pronunciation accuracy.

Much of Kendrick's data are provided in a descriptive format, rather than quantitatively. The author summarizes most of her findings, which are recorded in detail in her master's thesis published in 1995. The small number of participants makes it difficult to generalize the results widely. On the other hand, the study was long-term (9 months), collected a variety of data and repeated measures, used a range of instructional methods, used a realistic teaching environment, and included a range of tests from spontaneous speech to controlled reading.

In a study of 24 advanced NNSs of German (students at a U.S. university), Moyer (1999) investigated correlations among age, instruction, and motivation and ultimate L2 phonological attainment. Though the ratings of NNSs' pronunciation did not yet overlap with ratings of NSs' pronunciation of German, significant positive correlations ($r = .42$, $p = .049$) were found between amount and type of instruction and better (more native-like) pronunciation. In this case, learners who received suprasegmental instruction had more native-like German pronunciation.

Gorsuch (2001) wished to determine the effectiveness of pronunciation instruction using a specific English pronunciation textbook (*Clear Speech*, Gilbert, 1993). The participants were 24 second-year Japanese EFL students at a private Japanese college. They received 38 hours of EFL classroom instruction on perception and production of English suprasegmentals. Students completed the same speaking and listening tests pre- and post-instruction. Gorsuch found that after instruction, scores obtained from Gilbert's tests showed students improved in perception of

the target pronunciation features, but not in their production of these features. The author proceeded to compare the test to the actual instructional materials and concluded that key differences in the two resulted in a test that was not reliable and valid and the results of which were not generalizable to students' actual production capabilities. The researcher's own observations indicated that improvement had occurred, but the test was unable to detect the positive change.

Couper (2003) describes action research he conducted in his 16-week ESL class in New Zealand. Instruction consisted of approximately 1 hour per week in the classroom plus 2 hours per week in self-study. Participants were 15 adult professionals of whom 10 were Mandarin or Cantonese. Though controls were missing from his study, he collected pre- and post-instruction data, using read sentences and free speech and found that learners' errors in phoneme, word stress, strong and weak forms, epenthesis and absence, joining sounds, and sentence stress noticeably decreased over the 16-week course. Strengths of his study were the detail provided regarding the course content and how it was taught and specific error counts for the pre/post-tests. Instruction included

> analysis and explanation of pronunciation; controlled classroom practice; slightly less controlled classroom practice as learners added their own examples; listening to and recognising different aspects of pronunciation, especially in the language lab; language lab work involving listening and repeating; critically listening to their own speech after recording it in the language lab, both after a model and independently. (p. 59)

Derwing and Rossiter (2003) studied a group of 48 adult ESL learners in a Canadian college, who received 20 minutes per day of pronunciation instruction (out of 20 hours per week), over a 12-week training period. One experimental group received instruction on segmentals; the second group on global features (suprasegmentals: word and phrase stress, intonation, rhythm, projection, speech rate); a control group did not receive pronunciation

instruction. The rest of their language instruction was skills-based following a communicative curriculum. Expert raters rated speech samples from one reading and one spontaneous production task (story-telling from pictures). They rated for comprehensibility, accentedness, and fluency. Following the training period, only the global group showed significant improvement, and only in comprehensibility and fluency. The authors attributed the lack of impact of the segmental training on learners' information processing demands. They suggest that because of the "strong motor component" (p. 13) of segmental production, more attentional resources were needed and thus global pronunciation features did not improve for the segmental training group (though that group made significantly fewer phonological errors from time 1 to time 2, indicating that the instruction was successful for its intended goal). Because the global group received improved comprehensibility and fluency ratings the authors interpreted this as evidence for emphasizing the instruction of suprasegmentals over segmentals. However, this conclusion is likely not appropriate for learners whose sound-level accuracy is very low.

One possible limitation of the study can be noted: The raters evaluated prosody holistically. For segmentals, they could easily note errors on the transcripts they were provided, but notation of specific prosody errors was not included. My concern is that an overall rating might be different from a rating where specific errors were identified (location of PPS, fall/rise in intonation, word stress, MU boundaries).

Aufderhaar (2004) investigated ESL students' use of authentic audio materials (poetry, radio theater, short stories) in helping them improve their perception and production of suprasegmentals in English. She focused on Rost's (1990) Metrical Template Theory, which describes how language learners develop an internal "schema…for the prosody of the target language" (p. 736). Aufderhaar pointed out that students may start out filtering English speech

through the metrical template of their L1 and must develop an accurate schema for an English metrical template. Students were trained to listen to and analyze recordings and then performed them in class. Based on an 8-week instruction period involving eight students (including seven from China), Aufderhaar reported that the students' post-test ratios of stressed to unstressed vowels did not differ significantly from NSs', indicating that the intervention was successful in helping learners improve their production of English word stress. Raters' subjective evaluations of pronunciation and fluency suggested improvement, as did learners' self-reports.

Akita (2005) divided 64 first-year Japanese university EFL students into three groups, all of whom received the same basic instruction, except for the following: the *segmental* group (n = 23) received sound-level training; the *prosody* group (n = 24) received instruction on syllable structure, stress, and reduction; and a control group (n = 17) participated in extra conversation exercises in the place of segmental or prosody training. After 4 months of instruction (90 minutes x 12 sessions), all three groups improved in their perceptual abilities. Both experimental groups (but not the control group) improved their production of English prosodic features, and the prosody group improved the most on all measures, both segmental and prosody. Only the researcher rated speech samples one time, so reliability of ratings was not calculated.

Seferolug (2005) looked at two senior-year classes of 20 students each, preparing to be EFL teachers at a Turkish university. One group received 3 weeks of individual training using accent-reduction software (in a computer lab) and the control group received regular classroom pronunciation instruction. Both groups received the same amount of pronunciation training on the same sound, word, and phrase-level pronunciation features. Activities included perception, viewing waveforms of sounds, comparing waveforms of their own voices to that of a NS, viewing how sounds are articulated, practicing sounds, words, or sentences, and comparing to

NS. By the end of the 3 weeks, the control group's scores went down slightly, and the experimental group's scores were significantly higher. The study is limited in its generalizability to other proficiency groups and, like most studies in this review, does not show persistence of learning. However, the study is useful for demonstrating the use of computer-assisted pronunciation instruction in the language classroom and the results suggest that use of such software may be beneficial in an EFL setting where access to English is limited.

Effectiveness of pronunciation instruction for word stress. During two different sections of a semester-long oral communication course at a U.S. university, Murphy (2004) surveyed a total of 36 intermediate students' reactions to the use of a type of word-stress instruction. In his classes, Murphy used a numeric system in which students identified the number of syllables in a word and which syllable was stressed; this information was paired with gestures or movements timed with the major stress of a word. The majority (86%) found the system to be useful; more than 75% thought that the numeric system helped them pronounce new words and remember new words; and fewer (63%) said the system helped them use the new words in conversation. The author concluded that the instruction was helpful for this group of students, but expressed concern that 27% found the system too difficult. He suggests additional research is needed, for example, on what is a manageable number of word-stress rules or patterns for English for Specific Purposes learners at different proficiency levels; which patterns should be part of a core of instruction; comparisons of this technique to others and determining relative effectiveness; and qualitative investigations of learners' use of these patterns. Though the study does provide support for a specific technique for teaching word stress, unfortunately the author did not provide empirical evidence of improvement in his students' word-stress production.

Effectiveness of pronunciation instruction for PPS. The motivation underlying M. K. Hahn's (2002) research mirrored Kendrick's: She found no literature on the long-term effects of pronunciation instruction. This lack of data motivated her to investigate whether learning from an ESL pronunciation course at an American university persisted in the semesters following completion of the course. Hahn identified 36 NNSs of English who had taken the same ESL pronunciation class between Fall 1990 and Spring 2000 who were still on campus. They all showed substantial pronunciation improvement after one semester of instruction (T1 > T2). They were tested again from 3 to 13 semesters later to see if their performance remained above T1 testing. The results demonstrated that as a group, they showed persistence of learning of the nine PPS patterns studied in the original course. Individually 28 of 36 showed persistence of learning at T3; 27 of the 28 learners showed T3 scores that were lower than T2, but higher than T1. Only one learner showed performance where the persistence trend was T3 > T2 > T1. A limitation of this study is that the tests consisted of controlled elicitation (learners read dialogues after having some time to prepare). Thus we cannot generalize that the students could use the PPS patterns in spontaneous speech. Additionally, several variables could not be controlled, such as amount of L2 use, L1 background, and other learner variables. Hahn concludes that the more students learn during instruction, the more they retain over time.

Effectiveness of pronunciation instruction for intonation. DeBot (1983) performed one of the early studies investigating the relative influence of audio-visual versus auditory feedback in English intonation learning. Sixty-three Dutch university EFL students were placed in three groups: control, audio-visual feedback, and auditory feedback. The two experimental groups received 15 minutes of intonation instruction, followed by 45 or 90 minutes of the experimental treatment. Audio-visual training was found to be effective, and the auditory-only

group's performance deteriorated. The researchers concluded that the results were due to the fact that the audio-visual learners received visual feedback about their production, whereas the auditory learners had to rely on their own judgments of their production. The researcher acknowledged that this study did not address persistence of learning (the experimental treatments were a one-time occurrence), nor did it compare the experimental training to the effectiveness of a good language teacher.

Pennington and Ellis (2000) tested the recognition memory of 30 advanced adult Cantonese speakers of English, using sentences that were identical in all ways except for prosody. In the first phase of the study, the participants could identify lexical differences in sentences, but not differences due to prosody. In the second phase, participants received brief training in identifying prosodic differences, but were able to detect prosodic differences in only one of four sentence types. The researchers conclude that part of the explanation for the poor performance is the difference in the function of prosody in Cantonese (a tone language) and in English (an intonation language). Cantonese speakers are not accustomed to using prosody at the phrase level for interpreting meaning. They also concluded that unless L2 speakers' attention is drawn to the functioning of prosody in the L2, they likely will not attend to it, especially when the role of prosody is so different in their L1. Thus training is necessary to help L2 learners, even at an advanced stage, to notice and interpret prosody.

Harris (2003) offered eight weekly sessions of 1 hour each to university ESL students, who met individually with a first-year graduate student in speech-language pathology. Twelve Chinese students were in the experimental group and 15 were placed in a control group. The study's focus was on PPS and intonation (as well as rate and volume). To test learners' use of intonation the clinician asked a question and the participant had to use the same question each

time, but the use of PPS and intonation in the answer should vary based on the clinician's prompt. During training, the clinician primarily modeled correct intonation patterns, the client imitated, and minimal feedback was provided. Learners also were instructed to monitor their use of intonation while at home. No improvement or change in accuracy of intonation use was found. The clinicians' lack of experience and the limited amount of training (1 hour per week for 8 weeks) likely contributed to the lack of improvement.

Ewing (2002) investigated the effects of oral reading (25 minutes per day for 5 weeks) on the intonation in spontaneous speech of four Mandarin-speaking women (all were advanced speakers functioning at college level in English). The texts used for reading were marked with MU breaks and intonation contours. The researcher modeled the texts for each participant, and the participant shadowed the researcher's reading, read solo, and then read in tandem with the researcher. The author does not indicate whether feedback was given during these modeling sessions. The author found no significant change in intonation over the 5-week period and suggests that using less proficient speakers might be more appropriate in detecting an effect. Additionally the ratings were subjective and holistic, using a 5-point scale, rather than analyzing each utterance to determine accuracy of each intonation pattern.

After 10 weeks of training, 10 upper-intermediate undergraduate Spanish EFL learners in Ramirez Verdugo's (2006) study improved significantly in their abilities to notice and discriminate various English intonation contours and their meanings. Ratings of their spoken performance and intelligibility also increased significantly. The control group (n = 10) did not show similar improvement. During training, participants read dialogues and compared them to a NS model. The learners then viewed pitch displays of their and the NS model's speech and compared the displays and listened to the audio as often as needed.

Effectiveness of pronunciation instruction for rhythm. Couper (2006) investigated the effectiveness of teaching 21 high-intermediate ESL learners (two thirds Mandarin or Cantonese L1; remaining non-East Asian L1s) how to eliminate epenthesis and retain final sounds that often are omitted inappropriately, features that have great impact on rhythm. Instructional practices included listening to and modeling NSs, listening to and analyzing their own and classmates' speech, and listen-and-repeat activities. Couper found no significant improvement in learners' perception of the target features of epenthesis and sound absence, but did find a significant decrease in speaking error rates on reading tasks from week 1 to week 12, from 14% to 7.5% ($p < .05$). No change occurred in the baseline group, who received no pronunciation instruction over the 12-week period and maintained an error rate of approximately 10.5% during the period of observation. As with his 2003 study, Couper offered detailed descriptions of his teaching and in this study, he added a baseline group to allow comparison of the experimental group's performance to a similar group who received no instruction. A primary limitation of the study is that learners were evaluated based on readings of texts, rather than on error rates in spontaneous production.

Summary and Future Research Needed on Effectiveness of Suprasegmental Instruction

As noted at the beginning of this section, only 10 of the 22 studies on the effectiveness of instructional practice for the targeted content used adequate observational control for allowing interpretation of the results. Of those 10, two found no effect of instruction (Harris, 2003, for modeling and imitation; Macdonald et al., 1994, for drills, interactive practice, language lab use). The instruction in the former study likely was ineffective due to the use of inexperienced instructors and an apparent lack of corrective feedback. The latter study involved interventions

that were much too short to show an effect. The eight that found a positive effect focused on mixed practices (M. G. Chang, 2006; Couper, 2006; Kendrick, 1997; Ramirez Verdugo, 2006), contrasted instruction on segmental versus suprasegmental features (Akita, 2005; Derwing & Rossiter, 2003), or used speech visualization software for teaching prosodic features (deBot, 1983; Seferolug, 2005). The primary conclusions that can be gleaned from this group of studies are that extended rather than limited pronunciation instruction is effective, instruction on suprasegmentals (rather than segmentals-only instruction) is specifically needed if one wants to see the greatest improvement in suprasegmental features, and speech visualization software can be effective for teaching prosodic features in English.

As noted by Derwing and Munro (2005), the research and pedagogical literature now offers a convincing argument that pronunciation instruction on suprasegmental features is of value (deBot, 1983; Derwing & Rossiter, 2003; M. K. Hahn, 2002; Ramirez Verdugo, 2006; Seferolug, 2005). However, additional research is needed that demonstrates long-term effects of pronunciation instruction. M. K. Hahn (2002), Sardegna (2009; described in Chapter 2), and Kendrick (1997) were the only studies I found that focused on long-term persistence of learning and all reported positive results.

Another gap relates to what we know about the relative importance of various pronunciation features and thus their importance for teaching. The research is pointing toward suprasegmental features as being most important for intelligibility, but research on the impact of different contexts, learner factors, and mix of interlocutors suggest that the picture may be more complicated than simply being able to say one pronunciation feature is more important than another.

Much of the research on pronunciation instruction does not specifically outline how content was taught. For example, Couper (2003) and M. G. Chang (2006) are quite thorough in describing how and what they each taught, but many studies refer to terms such as communicative instruction or drills or language lab study without explaining more specifically the type or amount of each type of instruction.

Lack of adequate control in many studies makes it impossible to determine if an effect was due to the target intervention. Much of the literature on instructional practices for improving learners' control of the targeted content is non-experimental with limited controls and rarely are repeated measures or a comparison group used. Though naturalistic research is valuable for its realism, generalizability, and ecological validity, the lack of control makes it difficult to make claims regarding instructional effectiveness. Instructional interventions often are quite brief, particularly in earlier studies (e.g., MacDonald et al., 1994). Even extended interventions sometimes do not show improvements in pronunciation accuracy (Ewing, 2002; Gorsuch, 2001; Harris, 2003). Thus more long-term and adequately controlled work is needed to determine what reasonable expectations are for instructional effectiveness and for learner acquisition of each of the targeted content features.

Very little research is directed at instructional practices for specific pronunciation features or for specific instructional practices (Couper, 2006; M. K. Hahn, 2002; Murphy, 2004; Myers, 1995; Sardegna, 2009). This makes it difficult to be confident of the effects of a particular practice and map optimal practices to pronunciation targets. More research on instructional practice to target pronunciation feature mapping is needed.

Many studies offer evidence of improvement using global measures rather than using error counts that are categorized. For example, Couper (2003, 2006) provided specific error

counts, but Derwing and Munro (2003) provided only global measures for performance on prosodic features. The use of global measures makes it impossible to know which aspects of the learner's production changed or impacted listener ratings: Was it lack of durational contrasts in stressed and unstressed syllables? Was the pitch pattern inconsistent with the intended meaning? Was the PPS pitch jump on the wrong word? Future research is needed to better understand which pronunciation cues change following instructional intervention.

General and Pronunciation LLS Literature Review Summary

All LLS taxonomies identify some type of self-monitoring in their strategy lists. Learners frequently identify it as a strategy they use and researchers have found it to be one of the most commonly used by successful learners (Breun, 2001a; Reiss, 1985, cited in Peterson, 1997). Pronunciation texts and scholarly writers emphasize the need for learners to develop self-monitoring skills (W. B. Dickerson, 1984, 1989, 2000). However, several key gaps exist in what we know about the use of self-monitoring for pronunciation improvement: (a) We do not yet know which specific strategic behaviors are optimally included in self-monitoring. (b) We do not know how effectively learners can be trained to use such techniques and the role that proficiency may play in strategy use and effectiveness. (c) We do not know how effective certain self-monitoring techniques are for identifying and correcting specific L2 pronunciation features. This review has started to address the first two issues as follows: (a) By first situating within existing LLS and pronunciation strategy literature a model for the specific strategy clusters and activities that learners may be trained to use for monitoring and correcting their L2 pronunciation. This offers a model that can be tested with learners and with specific pronunciation targets. (b) This review offers a specific description of how learners have been or may be trained to use these

strategic behaviors. This also provides a model that can be used in studies of how learners may use the targeted strategies (listening, transcription, annotation, rehearsing) to monitor, evaluate, and correct their pronunciation of L2 suprasegmentals. This lays the groundwork for an appropriately controlled study of the effectiveness of self-monitoring strategies for learners' identification and correction of L2 suprasegmental features.

Research Questions

Given the findings of the literature review presented in this and the previous chapter, the following research questions were formed and are the basis of the current study:

1. Effectiveness of self-monitoring. (a) Does the use of self-monitoring, in general, enable learners to correct their nontargetlike pronunciation? (b) If so, do the three self-monitoring strategy types (L, LT, LTA) have differential effects on how much learners are able to make their pronunciation targetlike? (c) Do the three rounds of rehearsal (R) have differential effects on how much learners are able to make their pronunciation targetlike? (d) Does proficiency level relate to how effectively learners use the target strategies?

2. Correcting pronunciation targets. If the answer to 1(b) is affirmative, then do L, LT, LTA, and R have differential effects on how accurately learners orally correct each of the suprasegmental targets?

In the next chapter, I describe the teaching context for this study.

Chapter 4

Course Description: English Pronunciation for International Teaching Assistants

The speech data used in this study were collected during one semester of an ESL pronunciation course designed for international graduate students who needed to improve their English pronunciation and teaching skills prior to teaching in an English-speaking classroom. At the time of the current study, international students wishing to teach were required to pass the SPEAK test (Speaking Proficiency English Assessment Kit, published by Educational Testing Service, or ETS) in order to demonstrate adequate oral English proficiency. The SPEAK test is a retired version of the Test of Spoken English from ETS, a 20-minute oral test, and was offered once per semester in a computer lab. Students provided timed responses to 12 questions, and recordings of their responses were scored by anonymous raters. Rating was holistic, and pronunciation accuracy (intelligibility) was a significant consideration in rating. Students who received a failing score had to take an ESL oral skills course or work with a tutor prior to retaking the test.

Students who enrolled in the course typically fell into one of the following categories: students who were preparing to take the SPEAK test for the first time; those who failed the SPEAK test and had to take this course to qualify for re-taking the test; students who had passed the SPEAK test and wanted to improve their speaking, teaching, and/or pronunciation skills; and those who were not preparing to teach, but wanted to improve their oral English skills. For the 16 spaces available in the class, priority was given to current and prospective teaching assistants.

Placement and Diagnostic Testing

Any international graduate student could register for the class. However, on Day 1, a placement test was administered, during which students completed three unrehearsed tasks: reading a paragraph, describing a graph, and a 2-minute response to a prompt. Based on this test and the criteria described in the previous paragraph, students were assigned a priority for joining the class (if their greatest need was for pronunciation improvement) or a related course (if the greater need was for teaching and oral skills development). Class size was capped at 16, and if that limit was exceeded, a waiting list was created for lower priority students. On Day 2, enrolled students completed a diagnostic test which was similar to the placement test but contained different material. This test was used to identify learner needs, individualize instruction, and aid the instructor in prioritizing topic selection for the semester.

Course Principles

Key principles underlying the current study's course design included an emphasis on (a) increasing learners' awareness of the features of English pronunciation and how those features contribute to comprehensibility and intelligibility; (b) developing covert rehearsal skills, including self-monitoring strategy instruction; (c) identifying specific student needs; and (d) providing targeted pronunciation instruction and homework for reinforcing topics covered in class. Item b was specifically emphasized in the section of the course investigated in this study. Instructors of other sections varied in their emphasis on self-monitoring strategy instruction.

Pronunciation Instruction

Pronunciation topics included the sounds, rhythm, and melody of English. Suprasegmental instruction was based on two textbooks by L. D. Hahn and W. B. Dickerson (1999): *Speechcraft: Discourse Pronunciation for Advanced Learners* and *Speechcraft: Workbook for International TA Discourse.* Among suprasegmental topics covered were MUs, PPS, intonation, word and multiword construction stress, vowel reduction, alternating rhythm at the word and phrase level, and natural speech phenomena, such as h-elision, consonant cluster simplification (trimming), palatalization, and linking. Sound-level instruction was based on materials developed by L. J. Dickerson and Dickerson (1982, 1983) and the instructor's own materials. A general course outline appears in Table 5. A detailed version appears in Appendix D.

Table 5

General Course Outline for the Current Study

Week	Task	Comment
1 (1 class meeting)	Placement and diagnostic tests: text-reading and spontaneous production Completion of student information sheets	Used to determine students' placement in the class and individualized instruction needs
2	Minilecture 1	Audio recordings of this minilecture were used for assessing students' baseline use of the pronunciation targets and self-monitoring strategies
3-15	Classroom instruction on the target pronunciation features; training on self-monitoring skills	Instruction focused on perception, prediction, production of target features; training focused on development of self-monitoring strategies
10	Minilecture 2	Students practiced using the target pronunciation features and used self-monitoring
15	Minilecture 3	Audio recordings of this minilecture were used for the study's post-instruction tasks
16	Testing of strategy use Completion of post-task questionnaire	Comparison of use of the self-monitoring tasks in a computer lab, using tasks outlined in Chapter 5 and Appendix A

For all pronunciation topics, in-class time was spent helping students learn to perceive the targeted features by listening to NS models and to their own speech; learn rules for predicting the occurrence of these features; and gain experience and receive feedback on producing these features in a targetlike manner. Typical in-class and homework activities required students to listen to and identify a target feature, listen and repeat, predict features in prepared dialogues and then produce the dialogues orally, and participate in open activities requiring use of the feature in extemporaneous speech. Thus tasks were rotated, as appropriate, along a communicative continuum (Brown, 2007), among controlled, less controlled, and learner-constructed use of the target features.

Other homework and major assignments included the following: (a) Five audio recordings were assigned and contained four to six tasks at varying points along the communicative continuum. The assignments were focused on topics covered in class and were submitted online. The instructor provided individualized feedback to each student. (b) Written work was assigned to help students review, practice, and internalize in-class instruction. (c) Three 5-minute minilectures were assigned during the semester. The purpose was for learners to apply, in a mock teaching situation, the pronunciation skills and strategies covered so far in class. Students selected their own topics and used the vocabulary and style specific to their own academic disciplines. The minilecture was to be rehearsed prior to class, but students were expected to speak extemporaneously and were allowed to refer to notes or an outline. Few did so.

Self-Monitoring Instruction

Students were introduced to the self-monitoring strategies via classroom instruction and instructor-monitored practice in a computer lab setting. Homework assignments provided practice using the strategies with NS models and with students' own speech. Typical instructional activities included (a) listening to the instructor and repeating and using the key features in a targetlike manner (in class, in the lab, and during covert rehearsal); (b) during class, using a transcript to predict the location of target features, and then listening together to identify those features and discussing any inconsistencies; (c) transcribing a recording of a NS or the student's own speech, giving attention to pronunciation features covered in class; and (d) transcribing one's own minilecture recording, annotating, and audio recording rehearsals during which the student self-corrected nontargetlike productions.

Feedback

Feedback on pronunciation and strategy use was provided in three ways: (a) By the instructor, who provided immediate feedback during class-based activities; audio recorded individual feedback for each student's oral homework assignments; wrote comments on graded written homework; and met individually for 20 minutes with each student following each minilecture. (b) By peers: Immediately following each minilecture, students provided brief written feedback to three of their classmates, regarding comprehensibility of the speaker. (c) Self-evaluation by each student: Following each minilecture, students listened to their audio-recorded speech and completed a structured self-evaluation. This included a written portion containing overall impressions, transcriptions of selected portions, a detailed evaluation of their use of target features, and responses to peer and instructor feedback. Students had the opportunity to orally self-correct any nontarget features in a subsequent audio assignment.

Time Spent on Each Topic

The class met for 80 minutes, twice a week, over 16 weeks. Students did not attend class on three of the six minilecture days nor during the week of spring break. A total of 31 hours of face-to face instruction was offered. I estimate that students spent an additional 7 hours completing assigned homework. I cannot quantify amount of time spent in covert rehearsal or preparing for the three minilectures.

In Table 6 a breakdown of time spent on explicit instruction and homework is provided. During the semester, the following topics were covered, in descending order: strategies (12 hours or 31% of instruction); word stress (WS: 6.5 hours, 17%); sounds (5 hours, 13%); rhythm, including linking, vowel reduction, trimming, and blending (4 hours, 10.5%); primary phrase

stress (4 hours, 10.5%); construction stress (CS), including proper names of people, buildings, streets and compound nouns and numbers (3 hours, 8%); intonation (Int) (2 hours, 5%); and message units (2 hours, 5%).

Table 6

Topic Time Allotments for Explicit Instruction and Homework, in Descending Order

Time	Strategies	WS	Sounds	Rhythm	PPS	CS	Int	MU	Totals
%	31%	17%	13%	10.5%	10.5%	8%	5%	5%	100%
Hours	12	6.5	5	4	4	3	2	2	38

Note. Values are rounded to reflect the fact that these are estimates, not precise values.

Though I described and quantified the two types of instruction separately, instruction on strategies and pronunciation targets are interrelated, and difficult to separate. Use of the self-monitoring strategies requires use of the pronunciation targets and reinforces instruction on these targets. Both types of instruction were introduced simultaneously and learners were offered increasingly more demanding opportunities to practice the targets and strategies, until, by the end of the semester, they were functioning independently in their self-monitoring.

In the next chapter, I describe the study's methodology and how I investigated the effectiveness of self-monitoring strategy use for increasing students' accuracy on suprasegmental features.

Chapter 5

Methodology

As noted in Chapter 2, the use of the self-monitoring strategies for learning L2 suprasegmental features has received limited attention in the LLS and pronunciation pedagogy literature. An examination of the research literature reveals the need for a more systematic study to help us better understand whether the targeted strategies can be taught effectively to learners and the extent to which use of these activities results in improvements in L2 suprasegmental accuracy.

The purpose of this study was to help fill this need, by evaluating the effectiveness of adult L2 learners' use of three specific self-monitoring strategy combinations (the independent variable): (a) listening only + rehearsal (LR-LR-LR), (b) listening + transcription + rehearsal (LT-RRR), (c) listening + transcription + annotating (correcting) a transcript + rehearsal (LTA-RRR). The dependent variables were participants' accuracy scores, before and after using each strategy combination, for eight suprasegmental features: (a) message unit boundaries, (b) primary phrase stress, (c) intonation, (d) reduction of unstressed vowels in content words and (e) in function words, (f) linking, (g) word stress, (h) multiword construction stress. The study's context was described in the previous chapter. In this chapter, I present the methodology used to answer these research questions:

> 1. Effectiveness of self-monitoring. (a) Does the use of self-monitoring, in general, enable learners to correct their nontargetlike pronunciation? (b) If so, do the three self-monitoring strategy types (L, LT, LTA) have differential effects on how much learners are able to make their pronunciation targetlike? (c) Do the three rounds of rehearsal (R) have differential effects on how much learners are able to make their pronunciation targetlike? (d) Does proficiency level relate to how effectively learners use the target strategies?

2. Correcting pronunciation targets. If the answer to 1(b) is affirmative, then do L, LT, LTA, and R have differential effects on how accurately learners orally correct each of the suprasegmental targets?

Method

Participants. The original class, for which I was the instructor, had the maximum enrollment of 16 students. Fifteen students consented to their audio recordings being used in this study. From this point on, I refer to this group of 15 as the participants. Each participant is referred to by the letter P followed by a number representing their pre-instruction baseline proficiency ranking. For example, P1 refers to the participant with the highest pre-instruction proficiency rank and P15 is the participant with the lowest rank.

Permission to proceed with the study was received by the Institutional Review Board of the College of Education. Consent was obtained in the following way: Near the end of the second class session, I left the classroom while a colleague of mine handed out and then collected consent forms (Appendix E) from the students, in which they were asked permission for me to use in my study the audio recordings from their course assignments. The consent forms were held by the colleague, and the names of students giving consent remained anonymous to me until after I submitted final semester grades. All students in the class received the same instruction and completed all of the experimental tasks as part of required coursework.

As noted in Chapter 4, most students enrolled in the course were preparing to retake the SPEAK test or take it for the first time. Eight of the 15 participants had received a SPEAK score of 45 in a prior semester, 50 being the passing score. Three others received 40. One of the students who scored 45 (P9) in the previous semester received a passing score of 50 while enrolled in my class. Of the remaining four participants, three planned to take SPEAK for the first time in a subsequent semester, and one had taken and passed the test the previous semester.

Participants ranged in age from 23 to 35 years, 14 were L1 Mandarin speakers and one was an L1 Korean speaker. Time in the U.S. ranged from 9 to 33 months. Years of instruction ranged from 7 to 16. Seven students had prior pronunciation instruction. One student (P14) had taken pronunciation courses in China. P11 had five sessions with a pronunciation tutor. P9 had taken this same course one year previously; P1, P4, P10, and P12 had taken a related course in which pronunciation instruction was a minor component (primary focuses were culture, communication, classroom management, and lecturing skills).

Four students reported limited experience with self-monitoring: P1 had used listening, transcription, and rehearsal five or fewer times in a related ESL class; P4 used listening and rehearsal during three instructor office visits when taking the related ESL class; P7 used listening and rehearsal for TOEFL preparation and for practicing for a presentation; P8 used listening, transcription, and rehearsal five or fewer times prior to taking this course.

Demographic data and pronunciation and strategy-use background are summarized in Table 7.

Table 7

Demographic Data for the Seven Study Participants

Student	Sex	Age	Home country	Academic discipline	Months in U.S.	Prior pronunciation instruction?	SPEAK score	Prior strategy use?	Years of English instruction
P1	M	28	China	statistics	33	limited (UIUC)	45	yes	11
P2	M	26	China	computer science	10	no	45	no	13
P3	F	25	China	statistics	21	no	40	no	10
P4	M	25	China	biophysics	9	limited (UIUC)	45	yes	N.A.
P5	F	26	China	biology	9	no	45	no	10
P6	F	25	China	statistics	33	no	45	no	6
P7	F	24	China	engineering	9	no	40	yes	7
P8	M	35	Korea	architecture	33	no		yes	10
P9	F	24	China	sociology	9	repeated this course	50 (Spring 2009)	no	19
P10	F	27	China	industrial engineering	21	limited (UIUC)	50 (Fall 2008)	no	15
P11	F	24	China	biology	9	yes (limited tutoring)	45	no	10
P12	M	28	China	computer science	21	limited (UIUC)	40	no	11
P13	M	28	China	environmental engineering	21	no		no	15
P14	M	23	China	engineering	9	yes		no	11
P15	M	28	Taiwan	kinesiology	23	no	45	no	16
Mean		26.4			18				11.7

Materials. The speech samples used in this study were gathered from students' first and third minilectures. Each ML was approximately 5 minutes in length. MLs were presented during classes 4 and 5 and 27 and 28 (there were 29 class sessions during the semester). Students were instructed to explain a concept from their fields of study. The instructor approved students' topic choices and reviewed an outline prior to each minilecture. Students discussed their topics with classmates in the class period prior to presenting their ML. They were instructed to rehearse at home, but not to memorize the content, nor were they allowed to read from a written text during the ML. The last 1 to 2 minutes of each ML typically included the student's responses to audience questions.

Data from ML1 (pre-instruction) were used to establish a baseline of suprasegmental accuracy and strategy use. Data from the ML3 (post-instruction) were used to answer the research questions regarding effectiveness of strategy use for increasing suprasegmental accuracy. Each participant presented different content in ML1 and ML3.

Each minilecture was audio recorded in the classroom, using a Sony IC digital recorder (model ICD-P520) and a wireless microphone (Azden WM-Pro) attached to the student's collar. I roughly transcribed the minilectures in order to divide each one into six separate but equivalent audio files of 15 MUs. The first two files came from the first third of the ML, the next two files were from the middle third, and the last two files were from the final third of the ML. (Refer to Chapter 3 for a description of how MUs were defined.) These recordings were used by the students to complete the experimental tasks.

A questionnaire was administered following the completion of the ML3 self-monitoring tasks (see Appendix F). Numbered codes were used for each questionnaire to maintain participant confidentiality.

The procedures the students followed to complete the tasks are described in the next section.

Procedures

The procedures students followed when performing the self-monitoring tasks were the same for both MLs. Tasks for ML1 occurred during three consecutive class periods, but because class time for completing the tasks was limited, the ML3 experimental tasks were completed on one day during finals week, during a 2- to 2-1/2 hour session. To accommodate participants' final exam schedules, three participants completed the ML3 tasks on Monday, the others on Friday. For both ML1 and ML3, students were instructed to take breaks as needed following each set of tasks.

I did not track how long each participant took to complete each task. However, based on the length of recordings I prepared for each student, I estimate the following completion times: for L, 30 minutes; for LT, 25 to 30 minutes; and for LTA, 30 to 40 minutes.

Tasks were completed in the following order: listening + rehearsal (LR-LR-LR), listening + transcription + rehearsal (LT-RRR), and listening + transcription + transcription + rehearsal (LTA-RRR). The rationale for using this order was that later tasks built on earlier ones. Had some participants completed LTA first, there likely would have been carryover effects from using listening in LTA before using L, or carryover effects from using transcription in LTA before using it in LT. By having all participants use the same task order, carryover effects are not eliminated, but they should be the same for all participants.

For each participant, two of the six speech segments (an earlier one and a later one) were randomly assigned to each of the L, LT, and LTA tasks. This was done to prevent systematic

bias based on location of a segment in the original speech sample. However, for each experimental task, the speech segments were presented in the order in which the participant originally presented them, so that their production of the suprasegmental features would reflect the original discourse structure. For each student, I created a CD containing the task instructions (recorded in my voice) and audio files containing the segments from the student's ML.

The self-monitoring tasks were completed in a computer lab, with each student seated in front of a computer, with a headset that included headphones and a microphone positioned approximately one inch in front of the mouth. Each student was given the CD, written instructions that contained the same instructions as presented on the recording, and a checklist to follow as they listened. For the LT-RRR and LTA-RRR tasks, the written materials included a space for writing their transcriptions. See the complete instructions in Appendix A.

For each of the three tasks, participants inserted the CD in the computer. They opened a recording window in Audacity 1.2.6 (a free voice-recording program), started the CD recording in a program such as Windows Media Player, and followed the instructions. Next I describe subsequent steps separately for LR-LR-LR, LT-RRR, and LTA-RRR.

LR-LR-LR task procedures. Students listened five times to a 15-MU segment representing approximately one third of their lecture. They were instructed to listen for their use of one suprasegmental feature each time. The goal of this L phase was to self-monitor: listen and familiarize themselves with what they had said, and to give them time to focus on their accuracy on specific suprasegmental features. Following this focused listening phase, participants were presented with a 1- to 2-MU portion of the larger segment and were told to "Listen and repeat one time". This LR phase occurred three times for each portion until each 15-MU segment was completed. Based on findings from Bygate (2001) and Lynch and McLean (2001), my

assumption was that learners would be able to detect nontarget features following each rehearsal. I chose three rehearsals for pragmatic reasons: Going beyond three might result in fatigue or too great a cognitive load.

Participants recorded each rehearsal in Audacity. Following is an example of what one participant (P11) heard in her recording:

Researcher: "Time one, listen and repeat one time."

> Participant's ML recording: "Ok. so because it's a technique / so the most important thing is to know how it works and the procedure of this assay."
>
> [12 seconds of silence while the participant repeats, self-corrects, and records the same speech sample]

Researcher: "Time two, listen and repeat one time."

> Participant's ML recording: "Ok. so because it's a technique / so the most important thing is to know how it works and the procedure of this assay."
>
> [12 seconds of silence while the participant repeats, self-corrects, and records the same speech sample]

Researcher: "Time three, listen and repeat one time."

> Participant's ML recording: "Ok. so because it's a technique / so the most important thing is to know how it works and the procedure of this assay."
>
> [12 seconds of silence while the participant repeats, self-corrects, and records the same speech sample]

After completing the LR-LR-LR tasks for ML1, students used any remaining class time to complete other computer-based listening exercises provided by the instructor. After completing the ML3 L tasks, students were instructed to take a break and then return to complete LT tasks.

LT-RRR task procedures. Participants listened five times to a different portion of their lecture. During the first listening, participants transcribed the full segment. During subsequent

listenings, they focused separately on the targeted suprasegmental features and marked on the transcription what they heard themselves produce on the recording (Appendix A). After completing the Listening + Transcription, participants immediately read and orally corrected the transcribed segment a total of three times (RRR). Each participant audio recorded these three readings.

For ML1, students completed other listening tasks, as time permitted, as done for the L tasks. For ML3, students were instructed to take a break and then return to complete LTA tasks.

LTA-RRR task procedures. For the LTA task, participants completed the Listening + Transcription steps as described for the LT task (using a different section of the lecture). They then followed a checklist (Appendix A) and systematically reviewed the transcription for nontarget pronunciation and annotated corrections in a different colored pencil. Each participant then read and audio recorded the annotated transcription a total of three times (RRR). See Figure 2 for an example of an annotated transcription.

For ML1, students completed other listening tasks, as time permitted, as done for the L and LT tasks. For ML3, students left the computer lab after completing the LTA tasks and after I verified that their recordings were uploaded properly to the course website.

See Table 8 for a summary of the data collection procedures.

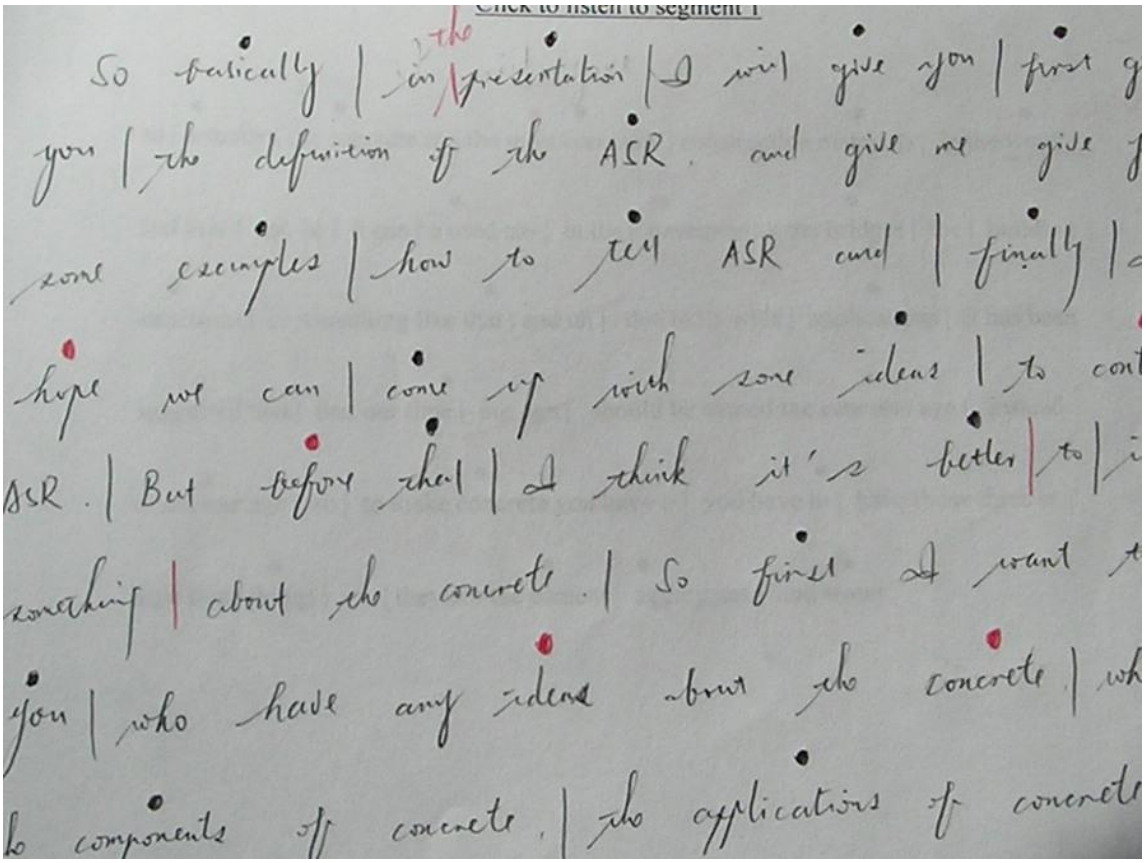


Figure 2. A learner's transcription (LTA) of a minilecture; annotations were marked in red.

Table 8

Summary of Data Collection Procedures.

1. Participants select ML topic, submit outline, and rehearse ML prior to presentation day.
2. Participants present their 5-minute MLs to the class (ML is audio recorded).
3. Researcher divides each ML recording into six segments; two segments (one early, one later) are randomly assigned to each strategy combination: L, LT, LTA.
4. Researcher prepares a CD for each student, for guiding the self-monitoring tasks. The ML content is unique to the student; instructions (Appendix A) are identical for all students.
5. In a computer lab, students complete the self-monitoring tasks and record themselves as they self-correct the original ML.
6. Researcher collects the recordings and prepares them for data analysis.

Note. Procedures are identical for ML1 and ML3.

Preparation of Data for Analysis

Because the original minilecture was recorded using a wireless microphone, the audio quality was lower than that of R1, R2, and R3. ML1 and ML3 recordings still were of sufficient quality for evaluating the targets. Only one modification was made to the audio recordings during transcription by the researcher: If a speaking rate was fast, I used the "Change Tempo" command in Audacity 1.2.6 to reduce the tempo by 10% to better perceive features such as linking and vowel reduction. This feature preserves the pitch and does not otherwise change or distort the recording. Following transcription of the audio recordings of each participant's minilecture and the three post-monitoring rehearsals, I marked the following eight features, according to what each speaker produced: MU boundaries, PPS, intonation patterns, word stress, multiword construction stress, vowel reduction in content words and in function words, and linking. Decisions about the presence of these features were based on the suprasegmental descriptions provided in Chapter 3. I also prepared a "target" version of the minilecture text (see Figure 3), which represented the suprasegmental features that, based on the semester's instruction, the students should be expected to produce accurately. This target version was the standard against which each ML and R1, R2, and R3 were compared, in order to produce the accuracy scores (dependent variables) used for the data analysis.

The following transcription conventions were used, according those used in L. D. Hahn and W. B. Dickerson (1999) and some of the conventions used in conversation analysis (Atkinson & Heritage, 1984):

Message unit boundaries were marked with a forward slash (/).

Primary phrase stress were marked using a solid black dot (●) above the syllable receiving PPS.

When a syllable that should be unstressed was given a heavy stress (but not PPS), an open circle (○) was placed above the syllable. This often occurred on pronouns that should be unstressed.

Intonation patterns were noted as follows: a comma (,) denoted non-final or question intonation; a period (.) denoted phrase-final intonation.

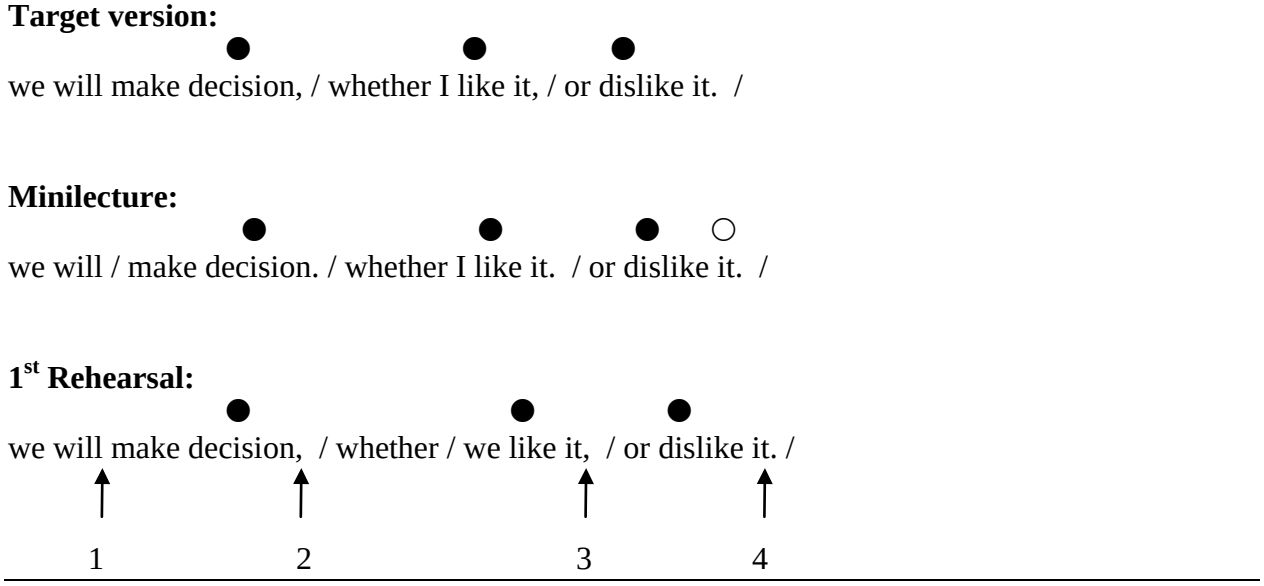


Figure 3. Data sample for P1: 1) deleting unnecessary MU break; 2) and 3) correcting intonation from final to non-final; and 4) de-stressing a function word (*it*) to highlight PPS.

Interrater reliability. A second rater completed an independent coding of a stratified random sampling of 10% of the data. Interrater reliability was 87.0%, using point-by-point agreement. The second rater's educational background was similar to the researcher's: completion of the MATESL degree, experience rating NS English speech, completion of a required phonology course, and at least one year teaching an intensive ESL pronunciation course. The researcher met with the rater to complete a 1-hour training session prior to the rating of the data and recalibration sessions were held twice.

Reliability ratings were highest for MU (95.3%) and Int (90.7%) and for the remaining three targets, agreement was somewhat lower:

MU	PPS	Int	FW	Link
95.3%	82.2%	90.7%	85.5%	85.3%
107 tokens	07 tokens	107 tokens	268 tokens	211 tokens

According to Kazdin (1982), agreement values of 80% or higher are regarded as acceptable. The lower agreement levels for PPS and Link are not that surprising. Decisions regarding prominence in a phrase can be subjective and determining the presence or absence of linking in fast speech can be difficult. Location of MU breaks, usually signaled by a pause, were much easier to detect.

Selection of equivalent speech samples. Message units were eliminated from the data analysis if the speech produced was not equivalent across all versions (i.e., minilecture and 1st, 2nd, and 3rd rehearsals). For example, if a participant added or deleted text in one version or revised the wording such that a particular MU was no longer parallel in content and phonological structure, the MU was deleted from the analysis. After unacceptable MUs were omitted, a total of 24 MUs per strategy combination could be used from each participant, for a total speech sample of 72 MUs per participant.

Accuracy and difference scores. Following the transcription of all the speech data and the elimination of unacceptable data, the researcher then coded each feature to indicate whether it was accurate, based on a comparison with what the student could be expected to do following the course's instruction (see earlier description of the "target" transcription). For each suprasegmental feature, the total number of correct targets were tallied, resulting in an accuracy score (represented as a percentage of total tokens possible for each targets) for each strategy combination used. For example, as shown in Table 9, during ML3, P11 produced PPS in a

targetlike manner in 10 of 24 message units, so her accuracy score was 42%. For the first rehearsal, her accuracy increased by two tokens, to 50%, and then on R2 and R3, her accuracy increased by seven tokens to 71%.

The ML accuracy score was considered the baseline level for each participant (what they could do without self-monitoring). Difference scores were calculated to compare each rehearsal to the baseline. These scores were used in the repeated measures analysis to determine the extent to which use of a specific strategy combination resulted in increased accuracy for each pronunciation target. Again, in P11's case, her difference scores were +8% for R1 and +29% for R2 and R3.

Table 9

Sample Data, Showing Computation of Accuracy and Difference Scores

P11, L only, PPS	ML3	R1	R2	R3
# of correct tokens out of 24 possible	10	12	17	17
Accuracy scores	0.42	0.50	0.71	0.71
		R1 – ML3 =	R2 – ML3 =	R3 – ML3 =
Difference scores		+0.08	+0.29	+0.29

Note. Accuracy and difference scores are in proportion form.

Accuracy and difference scores were calculated for each of the following data categories: (a) By strategy combination for all pronunciation features combined, resulting in scores for L, LT, LTA, and each rehearsal, for the total group and separately for each individual. (b) By suprasegmental feature, regardless of strategy type used, for the group and for each individual as well as for each rehearsal. c) By strategy type for each suprasegmental feature (strategy combination by target): for the group and for each individual and each rehearsal.

See Figure 4 for a summary of the study's procedures.

<table>
<tr><td colspan="6">1. Students present and record the ML during class.</td></tr>
<tr><td colspan="6">2. Researcher divides ML into six segments and randomly assigns each one to the L, LT, and LTA tasks (two segments per task).</td></tr>
<tr><td colspan="6">3. In a computer lab, participants complete one task for each segment:</td></tr>
<tr><td>Task 1:
Segment 1
(15 MUs)</td><td>Task 2:
Segment 2
(15 MUs)</td><td>Task 3:
Segment 3
(15 MUs)</td><td>Task 1:
Segment 4
(15 MUs)</td><td>Task 2:
Segment 5
(15 MUs)</td><td>Task 3:
Segment 6
(15 MUs)</td></tr>
<tr><td>L L L
↓ ↓ ↓</td><td>LT
↓</td><td>LTA
↓</td><td>L L L
↓ ↓ ↓</td><td>LT
↓</td><td>LTA
↓</td></tr>
<tr><td>R R R</td><td>R x 3</td><td>R x 3</td><td>R R R</td><td>R x 3</td><td>R x 3</td></tr>
<tr><td colspan="6">4. Researcher creates "target" transcription, to provide a baseline of accuracy and a total number of tokens for each pronunciation feature.</td></tr>
<tr><td colspan="6">5. Researcher creates transcriptions of ML3 and all rehearsals.</td></tr>
<tr><td colspan="6">6. Accuracy and difference scores are calculated.</td></tr>
</table>

Figure 4. Summary of study procedures. The same procedures were followed for ML1 and ML3.

Summary

In this chapter, I provided a description of the methodology used in the current study, which investigated the effects of 15 learners' use of self-monitoring strategies to increase pronunciation accuracy. I have described the participants and how the data were gathered and prepared for analysis. In Chapter 6, I present the results from the data analysis.

Chapter 6

Results

Two topics require discussion prior to reporting the results: (a) an evaluation of whether each pronunciation target had sufficient tokens for inclusion in the analysis, and (b) an interpretation of participants' pre-instruction use of the self-monitoring strategies. The study's results are presented after discussion of these two topics.

Selection of Post-Instruction Targets

Three of the eight original suprasegmental targets were eliminated from the pre- and post-instruction data analysis. Group accuracy gains were barely detectable for two targets for which participants had already achieved a high level of baseline accuracy: reduction of unstressed vowels in content words (CW: pre- and post-instruction baselines, 91% and 92%, respectively, with gains < 2.0) and word stress (WS: pre- and post- baselines: 95%; gains < 1%). And for construction stress (CS), participants produced few tokens overall (only 214 for ML1 and 269 for ML3, compared to 1080 each for MU, PPS, Int in both MLs, and 2245 for FW and 1878 for Link in ML3). CS use varied greatly, with six participants in ML1 and two in ML3 never using multiword constructions. Participants also made little improvement in accuracy for CS: less than 3% for ML1 and less than 4% for ML3. These limitations made it difficult to meaningfully interpret the results for CW and WS (due to ceiling effects), and CS (due to insufficient tokens and variable use across participants). Therefore, these three targets were omitted from the data analysis. Remaining analysis focuses on the following targets: message unit boundaries (MU), primary phrase stress (PPS), intonation (Int), vowel reduction in function words (FW), and linking (Link).

Pre-Instruction Use of Self-Monitoring

Prior to analyzing the full results of this study, I reviewed participants' pre-instruction accuracy scores (following the minilecture 1 self-monitoring tasks), to evaluate the effectiveness of participants' pre-instruction use of the self-monitoring strategies. I assumed that I would not observe significant differences in the accuracy scores for each of the three strategy types, because the participants had not yet received systematic strategy training in my course, and participants had not received prior training for these specific strategy types. And if this expectation were true, then pre-instruction data should represent the level of self-monitoring skill that participants had achieved as of the beginning of the semester. Data analysis could then focus on post-instruction tasks for answering the research questions: whether semester-long pronunciation self-monitoring training is effective. An analysis of pre-instruction data was needed to test this assumption.

Tests of within-subjects effects for Strategy Type confirmed that group accuracy scores were not significantly different among the three types[3]: $F(2, 28) = 2.527$, $MSE = .028$, $p = .098$. Group mean difference scores across all rehearsals and targets appear in Table 10. These findings indicate that, at the beginning of the semester, the different strategies did not have differential effects. This was the expected result: That without prior training, learner performance following strategy use would not be significantly different across the three strategy types.

I also had expected that prior to strategy training, participants might not be able to use the strategies with positive effect, given that their pronunciation knowledge and strategy experience were limited. To determine the size of the effect of strategy use on accuracy, I calculated

[3] The repeated measures ANOVA were calculated using SPSS, version 17.0.

Cohen's *d* (Kirk, 1995), using the following formula: ($\bar{x}_{R3}$ - $\bar{x}_{ML3}$) / s_{ML3}, or mean accuracy at R3 minus mean accuracy at baseline divided by the standard deviation of the baseline accuracy scores. The effect sizes for each strategy type ranged from medium to large[4], and for the targets, overall effect sizes were medium to large for MU, FW, Link, and PPS (Table 11). Effect sizes for Int were near zero. This suggests that even without specific training prior to instruction, the act of self-monitoring may lead to meaningful improvement for some learners and for some targets.

In order to answer the research questions, the remaining presentation and discussion of results focuses on post-instruction findings. I highlight interesting similarities between ML1 and ML3 as appropriate later in the presentation of the results and in Chapter 7.

Table 10

Strategy Effectiveness Following ML1, All Targets Combined

Strategy type	Mean difference score	Effect size (Cohen's *d*)
L	.021	0.5
LT	.050	0.6
LTA	.059	0.9

[4] According to Cohen (1988, cited in Kirk, 1995), d = 0.2 is small, 0.5 is medium, and 0.8 and above is large.

Table 11

Effect Sizes (Cohen's d) for Each Target Following ML1, by Strategy Type

	Targets				
Strategy type	MU	PPS	Int	FW	Link
L	0.8	-0.2	0	0.5	0.2
LT	0.8	0.4	0.1	0.9	0.3
LTA	0.7	0.6	0	0.4	0.4

In the following sections, the discussion is organized according to the research questions**:**

1. Effectiveness of self-monitoring. (a) Does the use of self-monitoring, in general, enable learners to correct their nontargetlike pronunciation? (b) If so, do the three self-monitoring strategy types (L, LT, LTA) have differential effects on how much learners are able to make their pronunciation targetlike? (c) Do the three rounds of rehearsal (R) have differential effects on how much learners are able to make their pronunciation targetlike? (d) Does proficiency level relate to how effectively learners use the target strategies?

2. Correcting pronunciation targets. If the answer to 1(b) is affirmative, then do L, LT, LTA, and R have differential effects on how accurately learners orally correct each of the suprasegmental targets?

Research Question 1. Effectiveness of Self-Monitoring

Group results. The answer to research question 1(a) was affirmative: In general, the use of self-monitoring appeared to result in improved suprasegmental accuracy. The baseline accuracy for the group, for all strategy types and targets, was 63.7%. The grand mean accuracy gain score was 7.0%, meaning that overall, participants' accuracy increased from 63.7% to 70.7% following self-monitoring. Effect size (Cohen's *d*) was 1.6, which is considered a very large effect size and much larger than that found for ML1 (0.9).

Strategy type. The answer to research question 1(b) also was affirmative: Two of the three self-monitoring strategy types had significantly different effects on how much learners were able to make their pronunciation more targetlike. Within-subjects tests showed a significant main effect for Strategy Type: $F(2, 28) = 4.867$, $MSE = .018$, $p = .015$. Pairwise comparisons of mean difference scores were significant at the $p = .05$ level for the L and LTA strategy types (Table 12). Other pairwise comparisons were not significant. Therefore, group accuracy gain scores following use of LTA were significantly greater than L accuracy gain scores.

Table 12

Pairwise Comparisons of the Three Strategy Types

				95% Confidence Interval for Difference[a]	
Pairs	Mean Difference	Std. Error	Sig.[a]	Lower Bound	Upper Bound
L to LT	.024	.012	.226	-.010	.057
LT to LTA	.016	.012	.645	-.017	.048
L to LTA	.039*	.014	.037	.002	.076

[a] Adjustment for multiple comparisons: Bonferroni.

*$p = .05$.

At the group level, LT is not significantly different from L or LTA. However, when looking at effect sizes, LT's is greatest, LTA slightly less so, and L's effect size is about half that of the other two:

Strategy	Mean difference score	Effect size (Cohen's *d*)
L	.045	0.78
LT	.069	1.6
LTA	.084	1.24

Interpretation of rehearsals data. Research question 1(c) also was answered affirmatively: Two of the three rounds of rehearsal had differential effects on how much learner accuracy changed. Results from within-subjects tests showed a significant main effect for Rehearsals, $F(1.485, 20.794) = 12.922$, $MSE = .004$, $p = .001$, with R2 > R1 and R3 > R1. Accuracy peaked at R2 and leveled off at R3. Both were significantly higher than R1 (Figure 5, Table 13). Effect sizes also provide a useful insight into the size of accuracy gains: all three Rs resulted in large effect sizes:

Rehearsal	Effect size (Cohen's *d*)
R1	1.2
R2	1.8
R3	1.7

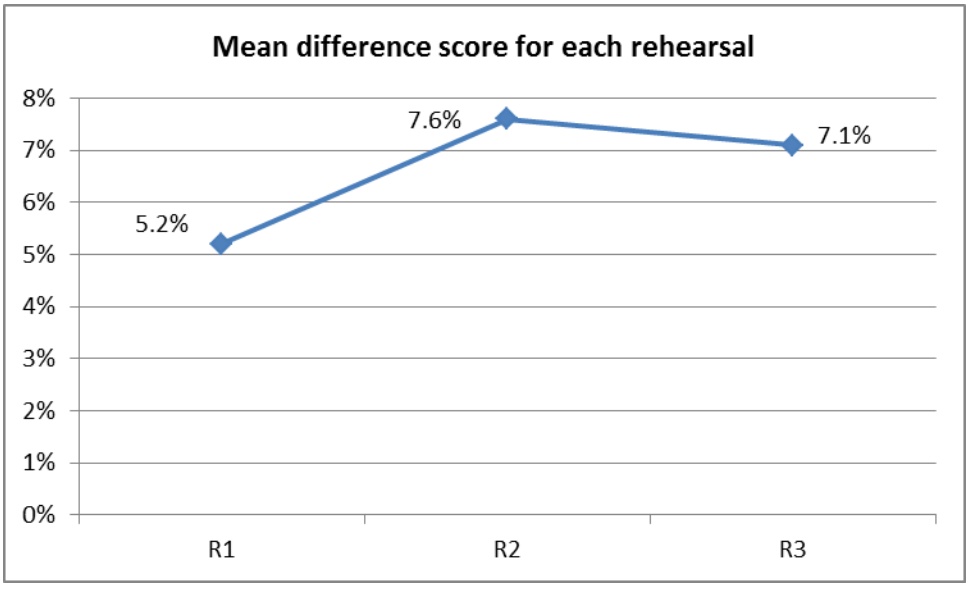


Figure 5. Mean difference score, by rehearsal, across all participants, strategy types, and targets. Accuracy peaked at R2 and leveled off at R3.

Table 13

Pairwise Comparisons, Across All Rehearsals and Targets

Pairs	Mean Difference	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
R1 to R3	.019*	.006	.027	.002	.036
R2 to R3	-.005	.004	.819	-.017	.007
R1 to R2	.024*	.004	.000	.013	.035

Note. Based on estimated marginal means

[a] Adjustment for multiple comparisons: Bonferroni.

*p = .05.

To simplify the presentation of data, the remaining analysis of the results focuses on difference scores at R3, because R3 represents the greatest average accuracy increase, across all participants, following use of self-monitoring. Performance at R1 and R2 will be discussed in sections relating specifically to the effects of rehearsal.

Proficiency level and strategy type. In order to find out whether the size of accuracy gains was related to baseline proficiency, I computed Pearson correlations for accuracy scores for ML3 baseline and R3 as well as for ML3 baseline and the difference score. A significant positive relationship was found for each strategy type between ML3 baseline and R3 accuracy (p = .05), suggesting that regardless of strategy type, accuracy gains tracked positively with baseline proficiency (Table 14). However, the size of the accuracy gain score had a negative but not statistically significant correlation with baseline proficiency (Figure 6). Thus lower proficiency participants appeared to achieve larger gains than did the higher proficiency learners, possibly because lower learners had more room to improve.

Table 14

Correlations for Each Strategy, Between Baseline Proficiency and R3 and Between Baseline and R3 Gain Score

Strategy	*r* (ML3Prof, R3 Prof)	*r* (ML3Prof, accuracy gain score)
L	.612*	-0.469 (*p* = .078)
LT	.578*	-0.451 (*p* = .091)
LTA	.676*	-0.253 (*p* = .364)

**p* = .05

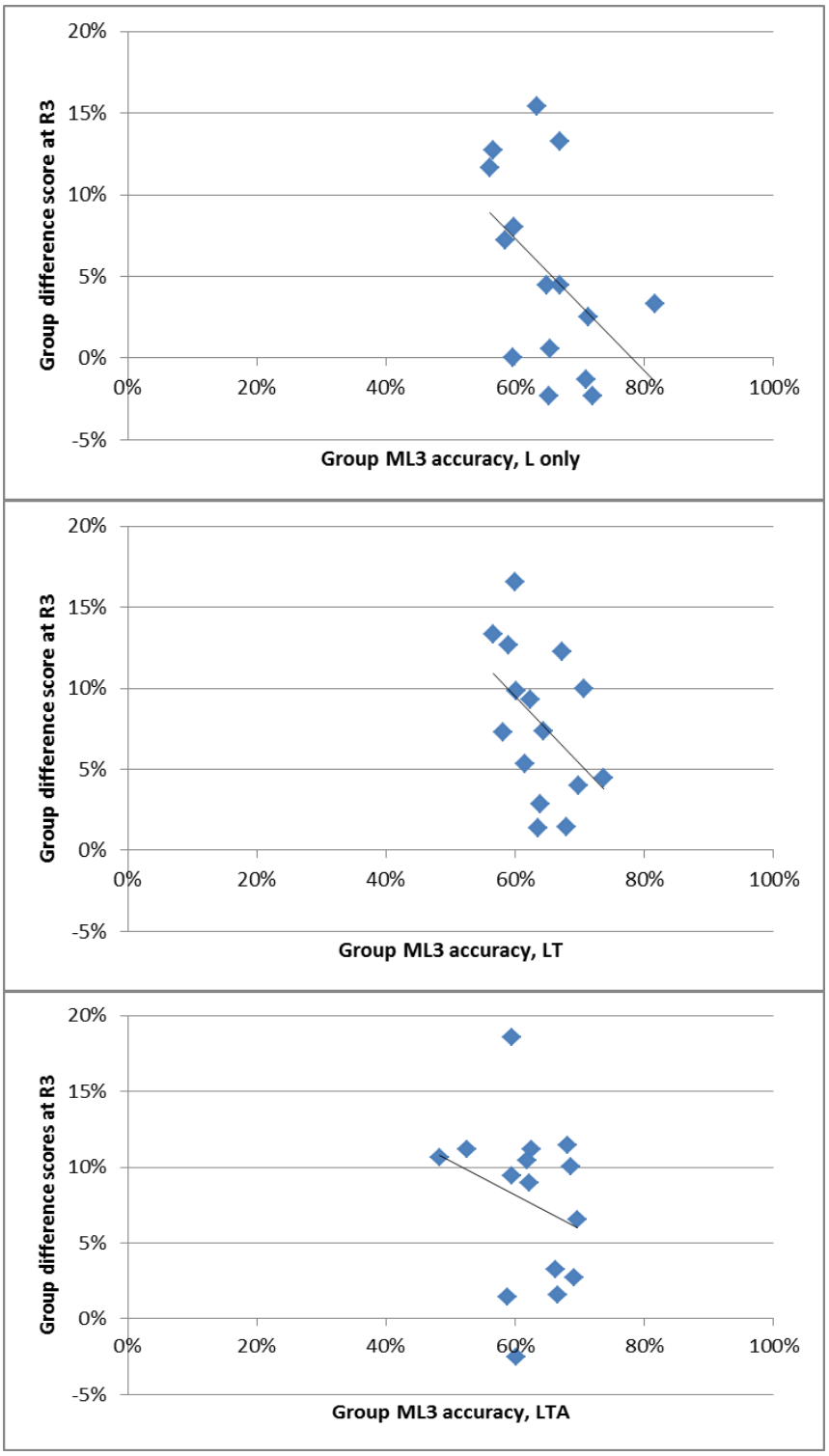


Figure 6. Scatterplots of group scores for ML3 accuracy (x axis) and difference scores at R3 (y axis), for L (top), LT (middle), and LTA (bottom). A negative relationship exists between these two measures: Participants with lower baseline accuracy scores tended to make larger accuracy gains.

Proficiency level and rehearsals. **A**ccuracy scores at R1 had a small positive relationship to starting proficiency (Table 15). However at R2 and R3, the strength of the relationship became negative, and lower proficiency learners' gain scores tended to be larger than those with higher proficiency.

Table 15

Correlations Between ML3 % Accuracy and Difference Scores for Each Rehearsal

R1	R2	R3
r = .075 (p = .789)	r = -.168 (p = .549)	r = -.286 (p = .302)

Individual results.

Overall performance. In Table 16, I present information about how each participant performed overall, across all strategy types and for all targets. The data are ranked according to the size of the difference score at R3. Also provided are values for ML3 % accuracy and R3 % accuracy, to help put the difference scores in perspective. Participant numbers also represent their overall ML1 proficiency ranking. The Pearson correlation for ML3 % accuracy and R3 difference score is small and negative: r = -.09. The correlation between ML3 and R3 accuracy is very strong, at r = .651 (p =.01), indicating that increasing accuracy scores tracked with rising proficiency scores (Figure 7).

Table 16

Participant Data in ML1 Baseline Order, Showing ML3 and R3 Accuracy and R3 Difference Scores, for All Strategy Types and Targets

Participant	ML3 accuracy	R3 accuracy	R3 – ML3 difference
P10	59.6%	72.7%	13.1%
P5	61.4%	74.2%	12.9%
P12	66.3%	78.8%	12.5%
P7	63.7%	72.5%	8.8%
P14	59.5%	68.1%	8.6%
P2	66.0%	73.8%	7.8%
P6	60.5%	68.3%	7.8%
P15	56.1%	63.8%	7.7%
P4	69.1%	76.6%	7.6%
P3	67.9%	72.8%	4.9%
P11	63.5%	67.6%	4.1%
P9	66.3%	70.3%	4.0%
P1	72.5%	74.9%	2.5%
P8	62.6%	64.3%	1.7%
P13	61.2%	61.8%	0.6%

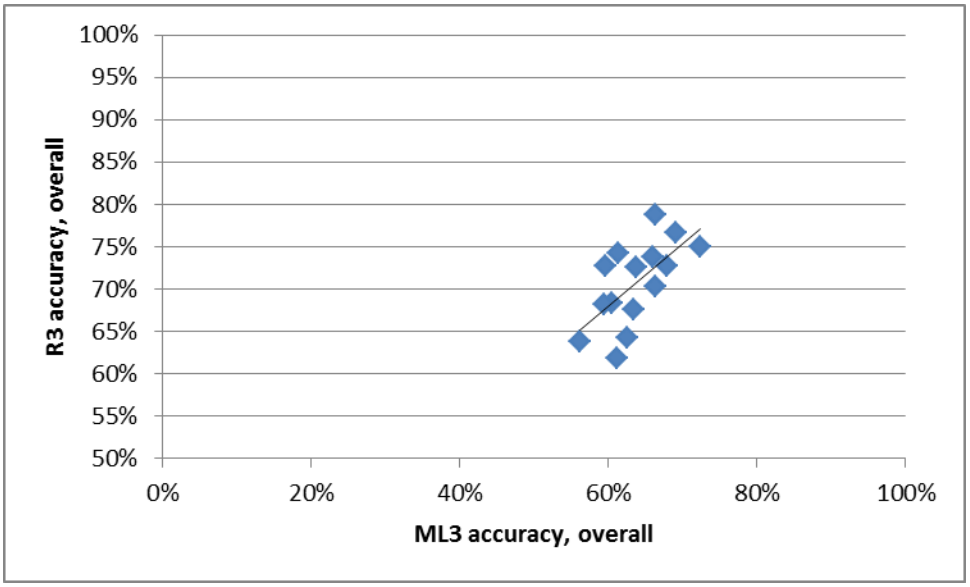


Figure 7. Scatterplot showing ML3 and R3 accuracy scores, for all strategies and targets combined. A positive relationship exists between these two measures.

Individual results, by strategy. In this section, I summarize difference score results at R3 according to individual performance for each strategy type. The results represent combined values for the five pronunciation targets. Separate results for each target are discussed in a later section.

In Table 17, participants are listed according to the strategy type that resulted in the greatest improvement. Use of LTA resulted in the largest accuracy increases for four participants, LT was most effective for seven others, and use of L was most effective for the remaining four. Ten participants were able to increase accuracy using all three strategy types. For four others, two strategy types were effective, and for one participant, only one strategy type was effective.

Table 17

Participants' Difference scores, by Strategy Type

ML3 rank	Participant	L	LT	LTA
10	P5	**15.4%**	12.7%	10.4%
7	P7	**12.7%**	7.4%	6.5%
8	P11	**8.0%**	2.8%	1.6%
1	P1	**3.3%**	1.5%	2.7%
13	P10	11.6%	**16.6%**	11.2%
12	P6	0.6%	**13.3%**	9.4%
2	P4	-2.3%	**12.2%**	11.4%
6	P2	4.5%	**10.0%**	9.0%
5	P9	-1.3%	**9.3%**	3.3%
11	P13	-2.3%	**7.3%**	-2.6%
9	P8	0.0%	**4.0%**	1.4%
4	P12	13.3%	4.5%	**18.6%**
14	P14	4.5%	9.8%	**11.2%**
15	P15	7.2%	5.3%	**10.7%**
3	P3	2.5%	1.4%	**10.1%**

Note. Each participant's greatest difference score appears in boldface.

Research Question 2. Correcting Specific Pronunciation Targets

Group results.

By target. Average difference scores for each target and across all participants and strategy types appear in Table 18. Group increases in accuracy were greatest for MUs (16.1%), followed in descending order by Link (7.4%), FW (6.1%), PPS (3.7%), and Int (2.1%). Effect sizes were large for MU, FW, and Link. Small effect sizes were observed for PPS and Int. Thus learners had the greatest success improving accuracy on MU, Link, and FW, and lesser improvement on PPS and Int.

Table 18

Summary of Overall Group Improvement, by Target and Proficiency Level

Measure	MU	Link	FW	PPS	Int
Difference score	16.1%	7.4%	6.1%	3.7%	2.1%
Cohen's *d*	1.8	0.7	1.0	0.4	0.3

Strategy type by target. In this section, I discuss group differences in the effectiveness of strategy types for each target. Group data appear in Table 19 and Figure 8.

LTA resulted in increased accuracy across all targets. Overall it was most effective for MU and PPS. LT was most the most effective strategy type for FW, Link, and Int, and nearly matched LTA in effectiveness for MU. LT resulted in no change for PPS. L never resulted in the greatest accuracy increase for a given target, though it resulted in sizeable accuracy increases for MU, Link, and FW, and a small increase for PPS. Group accuracy declined below baseline following use of L for Int.

Effect sizes were very large for MU for all strategy types, indicating that MU likely was an easy target for learners to improve. Link and FW also had medium to large effect sizes, and Int and PPS effect sizes were smaller.

Table 19

*Differences Scores and Effect Sizes (*Cohen's *d) for Each Strategy Type and Target (Group Data)*

Strategy type	MU	Link	FW	INT	PPS
			Difference score		
L	13.3%	6.4%	5.0%	-1.4%	1.9%
LT	17.2%	**8.1%**	**8.7%**	**5.3%**	0.0%
LTA	**17.8%**	7.7%	4.7%	2.5%	**9.2%**
			Effect size (Cohen's *d*)		
L	1.4	0.6	0.5	-0.1	0.1
LT	1.8	0.6	1.0	0.5	0.0
LTA	1.4	0.8	0.5	0.3	0.6

Note. Each target's dominant strategy is shown in boldface.

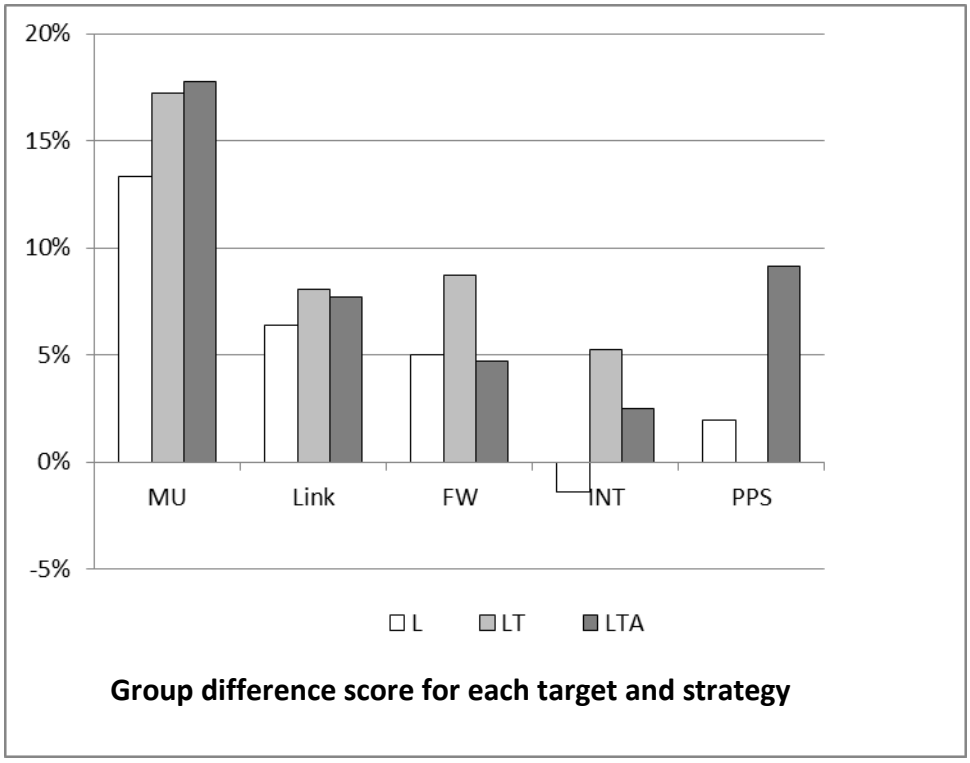


Figure 8. Average group difference scores, by target and strategy type. The strategies were similar in effectiveness for MU and Link. LT was best for FW and Int. LTA was best for PPS.

Individual results by strategy type and target.

Message units. Most participants made large accuracy increases for MU regardless of strategy type (Table 20), suggesting that MU was a relatively easy target for learners to monitor and correct. Ten participants reached post-monitoring accuracy levels above 90%, including four who achieved 100% accuracy. Correlations between ML3 accuracy for and difference score size for the three strategy types were negative and significant at the $p = .05$ level for L and at the $p = .01$ level for LT and LTA. This indicates that learners with lower baseline scores were making greater gains on MU accuracy following self-monitoring (Table 21).

An important point is that for MUs in the L only condition, participants did not need to identify the final MU break, because MU breaks were provided at the end of each 1- to 2-MU

chunk. How does this influence interpretation of the data? Participants did better using LT and LTA, suggesting that had I eliminated the final MU breaks from the analysis for L, participant accuracy scores might have been even lower.

Table 20

Difference Scores at R3, by Strategy Type and Participant (MU)

Participant	L	LT	LTA
P6	**20.8%**	8.3%	16.7%
P7	**16.7%**	4.2%	4.2%
P14	**29.2%**	**29.2%**	20.8%
P9	**12.5%**	**12.5%**	8.3%
P11	8.3%	**33.3%**	12.5%
P4	8.3%	**20.8%**	8.3%
P3	4.2%	**16.7%**	12.5%
P8	4.2%	**8.3%**	4.2%
P1	0.0%	**4.2%**	-4.2%
P5	4.2%	29.2%	**37.5%**
P10	16.7%	20.8%	**33.3%**
P12	25.0%	25.0%	**29.2%**
P15	25.0%	16.7%	**29.2%**
P2	4.2%	8.3%	**29.2%**
P13	20.8%	20.8%	**25.0%**

Note. Each participant's dominant strategy is shown in boldface. Participant number represents ML1 proficiency.

Table 21

Pearson Correlations, Comparing Overall ML3 Accuracy and Difference Scores for Each Strategy Type (MU)

L	LT	LTA
-.615*	-.775**	-.717**

*p = .05, two-tailed. ** p = .01, two-tailed.

Linking. Participants generally were successful at increasing Link accuracy following self-monitoring: Twelve participants increased accuracy with at least two of the strategy types (Table 22). P1 declined below baseline or remained unchanged. Given that he started with the highest baseline, this may represent a ceiling effect. Negative but statistically insignificant correlations existed between ML3 baseline accuracy for Link and size of difference scores at R3 (Table 23). As with MU, this suggests that lower proficiency learners were making somewhat larger gains for L and LT. The negative relationship was quite small for LTA.

Table 22

Difference Score Rankings, by Strategy Type and Participant (Link).

Participant	L	LT	LTA
P5	**36.4%**	10.5%	3.1%
P12	**18.6%**	0.0%	15.8%
P11	**12.0%**	-6.1%	-3.3%
P15	**11.8%**	4.3%	2.1%
P10	**9.7%**	9.4%	-11.4%
P6	4.4%	**29.5%**	15.9%
P9	-5.6%	**23.5%**	5.4%
P13	-4.2%	**21.1%**	-1.8%
P4	2.2%	15.7%	**25.0%**
P3	8.8%	-2.9%	**16.0%**
P2	6.7%	10.3%	**13.2%**
P7	4.8%	-9.1%	**11.1%**
P8	-11.4%	3.6%	**11.8%**
P14	2.6%	9.5%	**10.9%**
P1	0.0%	-6.9%	-2.8%

Note. Each participant's dominant strategy is shown in boldface.

Table 23

Pearson Correlations, Comparing Overall ML3 Accuracy and Difference Scores for Each Strategy Type (Link)

L	LT	LTA
-.343	-.582*	-.316

* $p = .05$, two-tailed.

Function words. Thirteen participants increased FW accuracy following use of at least two types of self-monitoring (Table 24). LT resulted in the greatest increases overall. Negative correlations existed between ML3 baseline accuracy for FW and size of difference scores at R3, indicating that lower proficiency learners tended to make greater increases. For L, the correlation was significant at the $p = .01$ level (Table 25).

Table 24

Difference Score Rankings, by Strategy Type and Participant (FW)

Participant	L	LT	LTA
P5	**21.6%**	19.6%	7.5%
P1	**14.3%**	-2.8%	2.6%
P13	**5.7%**	1.8%	-3.0%
P10	16.3%	**24.4%**	4.4%
P7	18.6%	**21.3%**	7.8%
P14	-4.3%	**16.3%**	11.5%
P6	-7.8%	**15.8%**	9.3%
P2	5.0%	**14.3%**	2.2%
P8	9.4%	**11.5%**	-14.3%
P11	-3.1%	**7.1%**	-1.7%
P4	1.8%	**5.5%**	1.3%
P12	17.6%	2.4%	**22.2%**
P3	0.0%	2.4%	**17.0%**
P15	4.3%	0.0%	**6.9%**
P9	-15.9%	-3.3%	-5.4%

Note. Each participant's dominant strategy is shown in boldface.

Table 25

Pearson Correlations, Comparing Overall ML3 Accuracy and Difference Scores for Each Strategy Type (FW)

L	LT	LTA
-.566*	-.358	-.258
$p = .028$	$p = .190$	$p = .353$

*$p = .05$, two-tailed.

***PPS*.** This target appeared difficult for some learners to improve (Table 26). Only three were able to improve using all strategy types. Two participants remained unchanged or declined and four others improved only when using one of the strategy types. When looking at proficiency, a strong negative relationship existed between starting proficiency and size of accuracy gain, meaning that lower proficiency learners were making the greater gains (Table 27).

Table 26

Difference Score Rankings, by Strategy Type and Participant (PPS).

Participant	L	LT	LTA
P11	**29.2%**	-8.3%	4.2%
P5	**8.3%**	-8.3%	8.3%
P7	12.5%	**16.7%**	8.3%
P15	0.0%	**16.7%**	**16.7%**
P1	4.2%	**12.5%**	**12.5%**
P2	0.0%	**8.3%**	0.0%
P10	12.5%	12.5%	**41.7%**
P14	12.5%	-4.2%	**20.8%**
P12	4.2%	-12.5%	**20.8%**
P4	-20.8%	0.0%	**16.7%**
P8	8.3%	-16.7%	**12.5%**
P9	-12.5%	0.0%	**8.3%**
P6	0.0%	-4.2%	**4.2%**
P3	0.0%	-12.5%	-16.7%
P13	-29.2%	0.0%	-20.8%

Note. Each participant's dominant strategy is shown in boldface.

Table 27

Pearson Correlations, Comparing Overall ML3 Accuracy and Difference Scores for Each Strategy Type (PPS)

L	LT	LTA
-.689**	-.558*	-.624*
p = .005	p = .031	p = .013

*p = .05, two-tailed. **p = .01, two-tailed.

***Intonation*.** Average baseline accuracy for Int was relatively high, at 83%, leaving less room for improvement. Strategy effectiveness was variable (Table 28). Three participants' accuracy remained unchanged or declined. Only two participants always improved, regardless of strategy type. Three participants reached 100% accuracy. A negative and highly significant correlation exists between baseline accuracy and gain size, indicating that lower proficiency learners were making the greater gains for Int (Table 29).

Table 28

Difference Score Rankings, by Strategy Type and Participant (Int)

Participant	L	LT	LTA
P9	**29.2%**	16.7%	8.3%
P7	**12.5%**	4.2%	-4.2%
P11	**8.3%**	-8.3%	8.3%
P2	**4.2%**	**4.2%**	**4.2%**
P4	-12.5%	**29.2%**	12.5%
P10	0.0%	**12.5%**	4.2%
P12	-8.3%	**12.5%**	4.2%
P8	-8.3%	**12.5%**	-4.2%
P5	-8.3%	**4.2%**	0.0%
P15	-4.2%	0.0%	**12.5%**
P3	0.0%	4.2%	**8.3%**
P1	-4.2%	4.2%	**8.3%**
P6	-8.3%	0.0%	-4.2%
P14	-8.3%	-8.3%	-8.3%
P13	-12.5%	-8.3%	-12.5%

Note. Each participant's dominant strategy is shown in boldface.

Table 29

Pearson Correlations, Comparing Overall ML3 Accuracy and Difference Scores for Each Strategy Type (Int)

L	LT	LTA
-.710**	-.443	-.716**
p = .003	p = .098	p = .003

**p = .01, two-tailed.

Targets by Rehearsals. Interaction effects on accuracy were found for Targets by Rehearsals, $F(12, 168) = 2.471$, $MSE = .317$, $p = .005$. For MU and FW, the accuracy ranking was R1 > R2 > R3. For PPS, Int, and Link, the ranking was R1 < R2 > R3 (Table 30). This pattern was the same for both proficiency groups.

Table 30

Difference Score Values for Each Rehearsal for Each Target

				95% Confidence Interval	
Targets	Rehearsals	Mean	Std. Error	Lower Bound	Upper Bound
MU	1	.127	.017	.089	.164
	2	.146	.019	.105	.188
	3	**.161**	.021	.116	.206
PPS	1	.019	.016	-.015	.053
	2	**.057**	.016	.023	.092
	3	.036	.024	-.016	.088
Int	1	.024	.015	-.007	.056
	2	**.040**	.018	.002	.078
	3	.021	.018	-.017	.060
FW	1	.034	.016	-.001	.069
	2	.050	.016	.016	.084
	3	**.064**	.017	.027	.101
Link	1	.054	.016	.019	.088
	2	**.086**	.020	.044	.129
	3	.071	.015	.038	.104

Note. For each target, the rehearsal with the greatest difference score is shown in boldface.

Summary

Research question 1 was answered in the affirmative: Self-monitoring use, in general, led to increased suprasegmental accuracy. Effect sizes for MUs (comparing accuracy at ML1 and R3) were large, suggesting learners may be able to detect the most salient targets without prior strategy training. Post-instruction gain scores and effect sizes were larger and may reflect the positive effects of strategy training. Use of rehearsal resulted in improved accuracy from R1 to

R2, but R2 and R3 were equal in amount of improvement achieved. The three strategy types had differential effects, with LTA being the most effective overall. Perhaps the more interesting finding relates to proficiency level and strategy type. In most cases, baseline proficiency was negatively correlated with gain score size, suggesting that lower proficiency learners were making larger gain scores than those at a higher proficiency level. Though such a finding might be expected, given that lower proficiency learners have more room to improve, this finding contradicts those of earlier studies (M. K. Hahn, 2002; Sardegna, 2009).

With regard to research question 2, self-monitoring appeared to have differential effects on suprasegmental accuracy. Participants were most successful at increasing accuracy for MU, Link, and FW and less successful for PPS and Int. At the target level, LT appeared to be most effective overall for Link, FW, and Int. LTA was most effective for MU and PPS.

An interaction effect was found for Targets by Rehearsals, suggesting some targets may be more difficult for learners to correct following repeated rehearsals.

In Chapter 7, I discuss these results, their pedagogical and future research implications, and limitations of the current study.

Chapter 7

Discussion of Results, Limitations, and Implications for Teaching and Future Research

The aim of this study was to investigate the effectiveness of training future ITAs to use self-monitoring strategies for correcting nontarget suprasegmental features in their own speech. The focus was on strategies that learners can use autonomously and with minimal reliance on specialized techniques or technology. In the following sections, I discuss how the current study's findings compare to earlier research on general and pronunciation LLSs; issues of proficiency and self-monitoring; and observations about what makes the target strategies more or less useful. Limitations, pedagogical implications, and suggested future research needs are discussed in the remaining sections.

How the Current Study Builds on Previous Research

General effectiveness of LLS use is confirmed. As noted in earlier chapters, research on strategy training has so far focused on two primary contexts: (a) Correlating learners' LLS use with their proficiency scores. These studies provide useful information, but do not tell us about the quality and flexibility of LLS use. (b) Some researchers have addressed this limitation and have provided, within an intact classroom, specific strategy training tied to L2 skills such as reading or writing. In these studies self-report data were used to identify strategy use. These studies still have the limitation that learners were not observed using a specific strategy at any given time. The current study addressed this issue by identifying four specific pronunciation learning strategies, training learners to use them over the course of a semester, and then observing learners' use of the strategies. This study satisfied the need for the ecological validity of the classroom setting, used repeated measures so that learners function as their own controls,

and provided evidence for strategy effectiveness for a majority of learners for one or more of the strategies and for five pronunciation features.

This study confirmed others' findings that LLS training can be highly effective for many L2 learners (Chamot & Kupper, 1989; Y. M. Chen, 2008; Cohen et al., 1995; W. B. Dickerson, 1989; Hassan et al., 2005; Sardegna, 2009). The current study also addresses some key limitations of earlier studies. First, participants were observed using a specific set of strategies designed for specific tasks. Other studies provided targeted strategy training, but left strategy selection up to the learners and thus strategy use was inferred but not directly observed (Chamot & Kupper, 1989; Cohen et al., 1995; Sardegna, 2009).

Though these latter three studies furthered our understanding of LLSs in important ways, and moved beyond what correlational studies could tell us, an intermediate step was missing: that of directly observing learners using clearly defined strategies that were matched to specific L2 tasks. Previous researchers emphasized that strategies should be matched appropriately to a given language learning task. But to date, I am not aware of other LLS studies that have observed when, how, and for which tasks learners used the target strategies. Thus the current study takes us another step toward filling this gap, by providing measures of specific strategy use paired with specific L2 features in a language-learning setting. I was able to measure strategy effectiveness by noting changes in target feature accuracy, pre- and post-strategy use. Though ultimately the goal is for learners to be able to select independently the optimal LLS to match the L2 task and their own individual set of learner variables (age, proficiency, learning preferences or styles), we still need to know, by direct observation, that a particular strategy indeed can be taught and that learners can deploy that strategy appropriately and gain the desired increase in L2 accuracy. Knowing that a strategy is effective allows us to confidently add it to the menu of

strategies we offer to language learners. The current study has demonstrated that this particular set of self-monitoring strategies indeed can be confidently added to the menu.

Feyten, Flaitz, and LaRocca's (1999) study on LLS awareness raising indicated that simply raising learner awareness about LLSs is not sufficient for increasing strategy use. The current study's findings are similar to those of Cohen et al. (1995): Learners may benefit from being shown how to use a strategy in a systematic way, and by doing so only one time, may be able to make meaningful accuracy gains. For example, following ML1, I gave learners checklists to follow (Appendix A) to guide their use of listening, transcription, annotation, and rehearsal. Interestingly, many learners made meaningful accuracy gains through self-monitoring, before they received instruction on the pronunciation targets and the strategies. And by ML3, effect sizes were nearly double those of ML1, indicating that sustained guided experience with strategies helps learners increase their strategic competence.

Pronunciation self-monitoring strategy training is worthwhile. At a more specific level, this study's findings confirm earlier research results on pronunciation strategy training effectiveness: that increased pronunciation accuracy results after training learners to use critical listening (Acton, 1984; Couper, 2003; W. B. Dickerson, 1987; M. K. Hahn, 2002; Sardegna, 2009), transcription (Acton, 1984; Mennim, 2003), annotation (Mennim, 2003); and rehearsal (Acton, 1984; Couper, 2003; W. B. Dickerson, 1987; M. K. Hahn, 2002; Sardegna, 2009). Only Mennim's study used annotation as a strategy component, but it was not the dominant focus. Thus the current study provides the first empirical evidence for the usefulness of the annotation phase as a strategy for further boosting pronunciation accuracy.

Additionally, transcription has not been highlighted as a specific pronunciation strategy. Only Eckstein (2007) lists it, but without description, in his taxonomy of pronunciation LLSs.

The results of the current study indicate that transcription of one's own L2 speech deserves a prominent spot in pronunciation LLS taxonomies.

Within this study, four students had used the strategies in a limited manner prior to enrolling in the course. This prior use did not appear to be related to how effectively these learners used the target strategies. Two participants with higher proficiency baselines had used some of the target strategies in a previous course. One achieved large accuracy increases following strategy use and the other did not. The first student's prior strategy use indeed may have helped her performance, or she was already a relatively competent user of LLSs and her higher baseline proficiency level may have assisted her as well. The other higher-proficiency learner did not achieve comparably high accuracy gains, suggesting that use of this particular set of strategies may not have been as effective for him. Ceiling effects due to his high baseline accuracy also may have been a factor. Two other participants had used some of the target strategies on their own prior to taking this course. One made large gains, which were consistent with her higher baseline proficiency level; and perhaps she was already an effective strategy user. The other made very small gains, which was consistent with his lower baseline proficiency. In his case, the strategies may not have been a good match for his learning preferences or he may not have been at the right stage of readiness to be able to self-monitor effectively.

Use of rehearsal is clarified. Rehearsal typically is taken for granted to be a useful form of language practice. Prevailing beliefs are that it may offer priming effects in working memory (Trofimovich & Gatbonton, 2006) that may enable focus on form, and provide an opportunity to notice nontarget production and correct it (Bygate, 2001; W. B. Dickerson, 1987; Lynch & McLean, 2001). I was unable to find prior research regarding the optimal number of rehearsals. The current results suggest that two are indeed useful. For the third rehearsal, learners usually

held even or declined slightly below R2 levels. And when accuracy declined at R3, participants typically reverted to an earlier incorrect form, though they sometimes made an error that did not occur in the original minilecture. This suggests several possibilities: Limits of working memory may make it hard to retain the desired corrected form, learners may revert to their own internal nontargetlike phonological models (J. Cole, personal communication, March 30, 2011), learners may be using a reading style, rather than using an oral style, and more training may be needed to help learners rehearse optimally.

Sometimes R3 resulted in the greatest gains, for example, for MU and FW. These are targets for which most participants made noticeable improvement. MU may be a more salient feature for learners, and reduction of vowels in FW may become easier with repetition: Perhaps through use of repetition, learners are able to increase their speaking rate and reductions may occur naturally. For PPS, Int, and Link, R2 was best. PPS seems, across the board, to be a more difficult feature to improve and learners may revert quickly to their internal phonological models. Lack of R3 improvement for Int could be due to ceiling effects. All three strategies are nearly equally effective for Link. Thus the strategy choice for Link may not be critical, rather learning to use rehearsal may be the appropriate focus for this target. Another possibility for why Link showed smaller gains is that learners may have felt they were maintaining clarity by pronouncing words separately. And in foreign language learning and speaking contexts and when the interlocutors are NNSs, lack of linking may not be an issue for intelligibility (Jenkins, 2002).

Yet another possibility is that the ability to correct certain pronunciation features may not be directly related to proficiency, learning preferences, or strategy use. Another factor, such as L1 interference, may influence learners' ability to perceive and produce certain features,

particularly when those features are used substantially differently in the L1 and L2 (see "Difficulties learners have in acquiring English prosody" in Chapter 3). This would be particularly true of stress and pitch features, such as PPS and Int, for the participants in the current study.

Implicit knowledge and second guessing. By consciously attending to L2 features that they previously produced automatically (implicit knowledge), participants may be "second guessing" themselves, that is, doubting their initial choices and changing something that was accurate to begin with to a nontargetlike form (Willingham & Goedert-Eschmann, 1999). If second guessing is a factor, one might expect greater declines on the more difficult, less transparent targets, such as PPS, FW, and Link. This possibility is supported by the observation that the target that appeared easiest for learners to change was MU, and R3 was indeed the most accurate rehearsal. But for PPS, Int, and Link, R2 was the best performance: second guessing could have impacted R3 accuracy. And this question of effects of second guessing may also relate to language aptitude, learning preferences, and other learner traits. Though outside the scope of the current study, these are interesting topics for future research.

Certainly the current study's preliminary findings are not sufficient to contradict long-held beliefs about the importance of rehearsal for improving L2 pronunciation. However, the findings do suggest that L2 learners may need additional training on how to maintain concentration on the learning task and thus maximize benefits from rehearsal.

New understanding of how strategy use affects suprasegmental accuracy. Provision of a detailed analysis of accuracy gains for five suprasegmental features is a key contribution of the current study to the pronunciation strategy and instruction literature. Most previous studies on pronunciation strategy use have not looked at accuracy gains by counting tokens of

nontargetlike productions made targetlike (Haslam, 2010) nor have the past studies that did use token counts to measure accuracy gains gone on to report the size of the gains (Acton, 1984; Couper, 2003; Mennim, 2003). Lack of such details makes it impossible to compare directly the results of those studies with my own. In general terms, the current study's results are consistent with these less detailed studies, in that, like the earlier studies, learners have been able to improve accuracy on targeted pronunciation features.

Three pronunciation strategy studies (W. B. Dickerson, 1987; M. K. Hahn, 2002; Sardegna, 2009) did analyze tokens of several stress-related features, but the data elicitation tasks involved reading a text, rather than observing spontaneously elicited speech. Like the current study, however, M. K. Hahn and Sardegna looked at post-instruction gains. In spite of differences in study design, the findings of the current study follows the same pattern found by these three studies: that of improved suprasegmental accuracy following strategy instruction and use. Because the current study looked at spontaneous speech, it provides an interesting comparison to the M. K. Hahn and Sardegna studies. The latter two found much larger gains in stress-related features than I did. This suggests that learners may gain greater suprasegmental accuracy in more controlled tasks before they are able to make similar gains in spontaneous speech. This is an important topic for further study.

Proficiency and Strategy Use

My findings relating to proficiency may help us understand the findings of others such as Haslam (2010), who found no relationship between reported strategy use and gains in foreign accent or fluency, and Sardegna (2009), who found that low proficiency learners' gains stayed lower than those of higher proficiency learners. Haslam's findings may be explained by the fact

that learners did not receive strategy training and were not observed using pronunciation strategies. Students were not directed to use a specific strategy for an appropriate target. This lack of strategy instruction may explain why strategy use did not have the expected positive effect on proficiency.

In the current study, the correlational measures indicate that lower level learners often made greater accuracy gains than learners that started at a higher proficiency, which is contrary to Sardegna's (2009) findings. Several explanations are possible. First, the groups of learners likely were not comparable. The learners in the Sardegna study typically were required to take a pronunciation class, according to placement tests given upon their arrival at the university. The students in the current study were exempted from that pronunciation class following the placement test. Thus the latter group's pronunciation proficiency likely was higher and they may have been behaving more like the High group in Sardegna's study. Another possibility may involve the nature of the strategies taught and how they were used by the students. The two strategic plans, covert rehearsal and self-monitoring, are different in some key ways and thus the outcomes of their use may differ. Sardegna's students used covert rehearsal outside the classroom and without instructor supervision. Covert rehearsal also does not involve self-recording and transcription. In my study, I observed the learners using the target strategies in the computer lab and through completion of homework tasks. The Sardegna study did not have such observational measures, other than what occurred during office visits. To better understand proficiency and pronunciation strategy use, we need research that addresses the inconsistencies in this small group of studies, by including direct observation of use of a variety of strategies and by learners at distinctly different proficiency levels.

Previous studies have observed a positive relationship between L2 learners' use of general and pronunciation LLSs and proficiency levels. The current study's results are consistent with these earlier findings. One refinement of the results of past studies is that learners starting at a lower entering proficiency make greater accuracy increases when using L or LT and that the higher proficiency learners are able to make large accuracy gains across all three strategy types. Additionally, the size of the accuracy gain appears to be negatively related to proficiency level. Like Sardegna (2009), I found that some learners who started out Low performed like High learners post-instruction, though she found this at T3, rather than immediately post-instruction. An important consideration in interpreting these findings is that a ceiling effect may have been in play for the higher learners.

What Makes the Strategies More or Less Useful?

Each strategy resulted in accuracy gains in most cases. Thus each of the strategy types has the potential to help learners focus on and improve suprasegmental accuracy. However, differences in strategy use emerged that appear, at least in part, related to individual differences.

Listening (L). L was the least effective strategy type overall, perhaps because it relies on one type of input (aural) and working memory limits may quickly be exceeded. However, the process of listening may allow the speaker to evaluate what he or she has produced and decide whether it "sounds right," i.e., compares favorably to an internal model of what is correct (W. B. Dickerson, 1989, 2000). When using L, nine participants made overall accuracy gains of 2.5% to 15.4%, indicating that for some learners the aural evaluation of their production was sufficient for correcting a meaningful percentage of nontarget suprasegmentals.

Interestingly, L was useful for all targets except Int. This observation is consistent with Pennington and Ellis's (2000) finding that Chinese L1 speakers have difficulty distinguishing intonation patterns in English when listening to sample phrases, most likely due to differences in how intonation is used in the L1 and L2. The extra steps of transcription and annotation may allow learners to focus specifically on pitch patterns and mark correct intonation patterns on the transcript, according to rules guiding intonation choice. These strategies may free up memory for attending to other details, particularly during rehearsal: The learner can look for the written cue to use rising, falling, or non-final intonation.

Another consideration, as mentioned earlier, is that all participants had achieved a relatively high level of accuracy for Int (> 80% on average), indicating that the remaining nontarget instances may be among the harder ones to correct. And yet a third possibility involving the LT and LTA tasks and use of rehearsal is that a task effect was in play: LT and LTA required reading a 24-MU segment, thus participants may have adopted a reading style, particularly later in the passage and by R3. Particularly for R3, I observed in the recorded speech data that some learners rushed through the rehearsals, particularly R3. This may have resulted in less accurate intonation.

It also may be easier for learners to hear MU errors, thus explaining why L is nearly as effective as LT and LTA for MU. Learners often report noticing lack of fluency, often reporting too many pauses, fillers, self-repairs. These are likely the most salient features and L may be sufficient for attending to them and correcting them.

Transcription (LT). Use of LT resulted in accuracy gains for all participants, ranging from 1.4% to 16.6%. Perhaps the act of transcribing slows down and focuses the critical listening process so the learner can identify the specific content and suprasegmental forms of his or her

utterances. By marking what was produced on the transcript, speakers can evaluate, as they transcribe, whether MU breaks, intonation patterns, pitch moves, alternating stress, and other features were produced accurately. Transcription also involves more varied forms of input (visual and kinesthetic) through the process of handwriting. These multiple inputs may reinforce each other and lessen cognitive demands. However, for PPS, the addition of Transcription had limited results. Only five participants improved PPS when using LT, indicating that this strategy type and PPS were not compatible.

A reason LT might be most effective for FW is because the simple act of reading the transcription may lead to more vowel reduction: Familiarity with the text may lead to more fluent production, as suggested by Bygate (2001) and Lynch and McLean (2001). The first reading may be more careful and subsequent ones more relaxed.

Annotation (LTA). Annotation offers learners the opportunity to apply pronunciation rules and revise their production. The positive effects of annotation in this study reinforce W. B. Dickerson's (1987) and Foster and Skehan's (1996) findings that learners were more accurate in correcting their own pronunciation after they monitored their own speech, rather than when they had time to apply rules (plan) before speaking. The annotation process may help learners focus on the form of their utterances, analyze their production and apply pronunciation rules, and then use the visual cues on the transcript to help remember to correct features when rehearsing.

Higher proficiency level may play a role for learners in the effectiveness of the LTA strategy type. Six of the eight lower proficiency participants made their greatest accuracy gains with L or LT. In contrast, five of seven higher participants made some of their greatest accuracy gains when using LTA. Several explanations are possible for this difference in use of annotation. First, the lower group may not have mastered the rules for predicting the five targets, making

annotation less useful and more likely to result in nontargetlike predictions. The higher learners likely had a greater mastery of the rules, and thus were more accurate when using rules during annotation. This finding is consistent with other studies that discovered that higher proficiency learners were able to use meta-cognitive strategies more effectively than lower learners (N. J. Anderson, 1991; Chamot & Kupper, 1989; Chesterfield & Chesterfield, 1985; O'Malley et al., 1985; Purpura, 1997).

A second option relates to how effectively learners use strategies. N. J. Anderson (1991) and Vann and Abraham (1990) found that less successful learners applied strategies inappropriately and lacked meta-cognitive or self-regulatory skills for evaluating a task and using strategies optimally. Again, this could be the case for the lower proficiency learners whose accuracy gain was very small or negative following use of LTA.

A third possibility is that annotation is necessary for certain learners to improve beyond their current proficiency level, or in other words, the use of annotation nudges them closer to 100% accuracy. Two lower learners and one higher made their greatest overall gains using LTA. Researchers have commented on variability in learners' ability to reflect on and regulate learning and performance (Ridley, 1997). Whether this is due to advanced readiness related to L2 proficiency, learning preferences that are well-matched to the target strategies, or increased abilities to self-reflect, is not yet established in the LLS research literature.

Why Do Some Learners Make Few or No Gains?

Though strategy use appeared to be effective for all participants at least for certain targets or strategy types, three learners made very small gains in overall accuracy. P13's gain was 0.6%, P8's was 1.7%; and P1, 2.5%. Most likely these gains were not statistically significant, and from

an instructional perspective, they likely were not meaningful. So why did the accuracy for these three learners remain essentially unchanged?

P1's baseline accuracy was the highest of the group, so his lack of progress may represent a ceiling effect. P8's baseline proficiency fell near the middle of the class's range and P13's was near the bottom. So for the last two, medium or small increases would be expected. Several other possibilities exist for explaining the lack of improvement. First, the strategies simply may not have been a good match for these students' learning styles or preferences. The target strategies required the user to be able to accurately hear his own production and these students' self-reflection or self-monitoring skills may not be strong. Ridley's (1997) study of reflection and strategy use in instructed L2 learning noted that language learners often vary in their preference or ability for self-reflection on L2 production tasks. She refers to various models of linguistic and learning behavior, including Bruner's (1960) continuum, with intuitive thinking at one extreme and analytic thinking at the other. Another relevant continuum described by Ellis (1992; cited in Ridley) focuses on a preference for accuracy versus fluency. Perhaps these three learners (P1, P8, and P13) were not analytic thinkers (at least in terms of pronunciation), or may have been more concerned with fluency than accuracy. Thus a number of variables related to learning preferences or styles or even the learning context itself may have affected the performance of these three learners in the current study.

Isaacs (2006) found that some of her lower proficiency participants also scored lower on measures of auditory aptitude (though the finding was not statistically significant), suggesting that some learners simply have more difficulty perceiving pronunciation features in their own and others' speech. Amount of L2 exposure also is frequently found to be a significant factor in L2 acquisition (Riney & Flege, 1998), with those living longer in an L2 environment and using

the L2 predominantly receiving lower ratings on accentedness. P1 and P8 had been in the U.S. for 33 months and P13 for 21 months. However, they reported their daily English use at 60 minutes (P1, P8) and 120 minutes (P13). Thus lack of English use outside the classroom may have had a negative effect on L2 pronunciation improvement.

LLS studies often use various types of instruments for identifying learning styles, preferences, and personality traits and relating these factors to strategy use. Having such data could shed additional light on these three learners' performance.

Clearly, the results of this study indicate that the target self-monitoring strategies do not result in increased accuracy for all learners. And past LLS researchers have identified in their studies small groups of learners who simply do not progress following language instruction (Breun, 2001b; Chesterfield & Chesterfield, 1985). This raises an important topic for future research: how best to assist learners for whom pronunciation progress is especially slow or difficult.

Limitations

In the preceding sections I noted several limitations of this study. Next I provide a summary of these and other limitations that must be considered when interpreting the study's results.

1. First, the sample size was small. The results are most generalizable to the teaching context within which the data were gathered: pronunciation courses for ITAs.
2. The fact that 14 of 15 participants were L1 speakers of Mandarin was a coincidence. Having such a homogenous group of learners certainly was useful for looking at what Chinese speakers of English may do with pronunciation strategies, and generalizing to this very large group of speakers may be useful and appropriate. The one Korean L1 speaker's (P8) accuracy gains were small, near the bottom of the group. However, the patterns of his accuracy score gains were not noticeably different from the Mandarin L1 participants. P8 was the oldest participant (age 33), which, as noted earlier, may have been the more relevant factor in his low accuracy gains. The homogenous sample in the

current study may not be a serious disadvantage, because most general and pronunciation strategy research has occurred in a variety of L1 and L2 contexts and those findings regarding strategy use have been predominately positive.

3. Even though I discussed the role of proficiency level in this study, the learners were homogenous, in the sense that all were at an advanced level of overall oral English proficiency, but still needed help with pronunciation. Whether the targeted strategy training would be similarly effective for learners at low and intermediate oral proficiency levels is not definitively known. However, studies cited earlier indicate that instructors should expect that lower proficiency learners may need explicit instruction in using meta-cognitive strategies.

4. Inconsistencies in task types may have influenced outcomes. For example, some tasks required reading, others listening only. Also, the original minilecture was presented to an audience, however, the experimental tasks were not. And though the focus of this study is on strategies for self-directed study, an interesting question is to what extent interaction may result in more accurate production when using self-monitoring. Ridley's (1997) learners varied in their preferences for learning in more analytic or communicative contexts, and finding ways to encourage language learners to develop strategic competence in communicative settings makes intuitive sense. Further work is needed to understand how such task differences may affect the study's results and impact strategy effectiveness.

5. With multiple readings in the LT and LTA tasks, learners may have started to rush and focus less on producing accurate target features during R3. The findings from the current study conflict somewhat with those of Lynch and McLean (2001), who found that repetition consistently resulted in pronunciation improvements, due to task familiarity. Further investigation is needed to better understand the current findings.

6. Memory effects: Learners may be over-taxing short-term and working memory as they process multiple chunks of text during rehearsal. In this study, I required learners to focus on eight suprasegmental features during the experimental tasks. Though the instructions guided them to attend to one target at a time in a systematic manner, during the rehearsal phase they were expected to correct as many of the eight targets as possible. Learners very likely chose, consciously or unconsciously, to focus on a subset of the eight features as they implemented the strategies. And their focus may have been on different features during each of the rehearsals. And as noted in Ridley's (1997) study, some learners may have been more focused on accuracy and others on fluency. In future research, this problem could be addressed by having learners use each strategy type with only one feature at a time.

7. The results do not offer insights into long-term effects of self-monitoring, though at least two studies give us some very convincing evidence of long-term benefits of covert rehearsal strategies (M. K. Hahn, 2002; Sardegna, 2009).

8. The tasks used in this study were not true examples of covert rehearsal. Rather than completing tasks in privacy, participants completed the tasks in a computer lab that

resembled a testing situation. However, in order to ensure tasks were completed in a consistent manner using equivalent speech samples, the artificiality of a semi-experimental setting was necessary.

9. Learners often avoid forms that are unfamiliar or have proven to be problematic for their intelligibility and LLS research has found that this is a strategy that learners consciously employ (Cohen et al., 1995). Because I used spontaneously produced speech in this study, I cannot claim that the speech samples represent learners' true pronunciation proficiency. For example, accuracy levels for word stress exceeded 95%, yet I know from working with this group of students that they frequently used inaccurate word stress in other contexts, particularly when asked to produce less familiar words. The same likely is true for vowel reduction in content words and intonation (learners demonstrated high baseline accuracy, but inaccuracies were observed in the classroom and in homework). Future researchers should continue to explore ways to elicit various types of speech (spontaneous and controlled) to gain a full understanding of learner proficiency and the effectiveness of pronunciation strategies on read versus spontaneous production.

10. Another item missing from the current study is an assessment of learners' pre-instruction pronunciation strategy use. I did determine that four learners had very limited prior experience with the target strategies, and that none had received focused training prior to taking this class. However, learners may have been using other strategies in combination with the target ones. Thus administration of a strategy survey (such as SILL, Oxford, 1990) at the beginning of the semester may have aided interpretation of this study's results.

11. Another important limitation is that I did not survey the students to determine the extent to which the target strategies might align with their individual learning preferences or styles. This was not within the scope of the current study, but the learners who did not perform as well may have had limited improvement due to these and other factors.

As the above list of limitations demonstrates, one must use caution in generalizing beyond the teaching context of this study. However, as noted earlier, this study does provide useful information about the effectiveness of training adult L2 learners to use a specific set of strategies for improving pronunciation accuracy. I hope these findings will allow others to take the story a few steps further as the limitations of this study are controlled for in future research.

Other Factors That May Affect Interpretation of the Results

Fatigue. Participants completed the tasks during one session of 2.5 hours or less. This was done to accommodate to the institutional constrains of the semester and course schedule:

The amount of class time available for completing the experimental tasks was limited. Ideally, the tasks should be completed in shorter sessions over several days, and in actual practice, this is indeed what L2 learners would do. LTA was always the final task, which would lead one to expect the greatest effects of fatigue when using this strategy combination. However, this did not occur: LTA was often the most effective strategy type. However, the increased accuracy using LTA could be due to practice effects, from monitoring the target features repeatedly for L and LT before getting to LTA. A way to address this limitation is to have learners perform the self-monitoring tasks on different days.

The design of the L task required participants to listen to their ML3 in small chunks of one or two MUs, following Miller's (1956) "magical number seven, plus or minus two" guideline (see Chapter 3, "Characteristics of targetlike English MU boundaries"). Keeping the chunks within this range decreased the likelihood of overloading working memory. Except for PPS, L was least effective for most learners, suggesting that participants may have felt fatigue, boredom, or frustration over the repetitive nature of the task. In fact, one participant voiced frustration in his recording regarding the slow pace of the task. Because L was a useful task for some participants, I would want to explore ways to redesign this task to make it more effective for learners who are more attuned to aural strategies.

Lack of motivation to complete the task accurately. Motivation levels of L2 learners likely vary. Eleven of the 15 students enrolled in the class were required to take the class in order to retake the SPEAK test and three others were there to prepare for the test. Some learners may have simply been fulfilling the requirement with the minimal amount of effort. Others may have believed participation in the class was important for passing the test. Motivation levels were not explored in this study. However, on a post-task questionnaire, the average ratings on a 5-point

scale (5 = very useful; 1 = not useful) for listening and rehearsal were 4.7 for each, but the rating for transcription was 4.0. I suspect that transcription received a lower rating due to the effort involved in using it. This suggests that learners may benefit from concrete evidence of how they performed on each of the three strategy tasks. Such evidence could help them better determine which strategy components were, in practice, most effective as well as convince them that the strategies indeed result in improved accuracy.

Another observation regarding motivation is that, from purely subjective observations of my own, the three lowest performing learners appeared to be highly motivated and diligent in completing the experimental tasks. Other course-related behaviors such as regular attendance and active participation, and completion of homework assignments also suggested these learners were motivated to complete the coursework and improve their pronunciation. Conversely, two learners who seemed to rush through the strategy tasks, attended irregularly, and participated less in class were ones who made larger increases in pronunciation accuracy. Thus something other than motivation may be at play in determining strategy effectiveness.

Pedagogical Implications

Descriptions of the pronunciation and strategy instructional methods and materials used during this study appear in Chapters 2, 3, 4, and Appendix A. I repeat here the objectives I used for designing the strategy instruction:

1. Students overcome self-consciousness when listening to their own speech and learn to be disinterested but critical listeners
2. Students become aware of the target pronunciation features and value of pronunciation strategies such as self-monitoring, self-evaluation, self-correction
3. Students are exposed via classroom activities to models of the target pronunciation features and strategies

4. Students practice using the pronunciation features and strategies in class, in homework assignments, and in contextualized activities such as minilectures and subsequent transcription and correction
5. Students learn to evaluate their production of targeted content and strategy use through written self-reflections
6. Students learn how the strategies can be used for other tasks such as monitoring other pronunciation and oral English skills

Most of the learners in this study achieved these objectives, but as noted previously, three learners did not make substantial accuracy gains following strategy use. Does this mean that this approach was not useful for these learners? In the end-of-semester questionnaire, these three rated the strategies according to "its role in helping you improve your English (5 = very useful; 1 = not useful)". These three participants rated the strategies at the 4 and 5 level. Thus, in some way, they saw the strategies as useful, even if objectively, their production did not change substantially overall. This is important to remember when observing learners' strategy use: Students may be benefitting in ways we cannot directly observe. However, substantial declines in accuracy following strategy use signal that the instructor must intervene and decide whether the strategies are a mismatch or whether the student simply needs to retreat a few steps and focus on earlier steps in the self-monitoring process.

Following are additional points for consideration when planning pronunciation strategy instruction.

Expectations for pronunciation improvement and strategy use. The current study suggests what is frequently observed in the language classroom: Accuracy in spontaneous production takes time to develop, at least for certain features. Studies looking at effects of longer-term suprasegmental instruction (8 to 16 weeks or more) typically reported pronunciation improvement (Akita, 2005; Couper, 2003; Derwing & Rossiter, 2003; M.K. Hahn, 2002; Kendrick, 1997; Myers, 1995; Sardegna, 2009). Studies looking at short-term interventions saw

limited or no effect (Ewing, 2002; Harris, 2003; Macdonald et al., 1994). Additionally, simply raising learners' awareness of LLSs (Feyten et al. 1999) does not lead to increased strategy use. Giving learners experience using the strategies is an essential step. Some learners may use the targeted pronunciation strategies effectively with minimal guidance, whereas other may need regular assistance and feedback as they gradually learn to use the strategies.

Prioritizing pronunciation instruction. In any pronunciation course, topics must be prioritized according to learner need, which can be determined via a diagnostic test given during the first week of the semester. Assuming that learners need assistance with the five suprasegmentals targeted in this study, the following principles can be used in setting priorities.

Learners seem to understand quickly the concept of MUs, seem to easily hear what they are doing with this target, and make rapid improvements. This suggests that this should be covered early. If learners make rapid progress, MU may not require additional classroom focus. (However, even if MU is not the focus of a lesson, I recycle pronunciation content throughout the semester by reminding learners regularly to attend to this and other features covered earlier in the semester.) Learners also may benefit from use of the strategies early on with initial focus only on MU boundaries.

Similarly, learners may be able to more easily grasp the rules for assigning PPS (new/old information; contrasts). However, the accuracy gains following self-monitoring are smaller than for MU, at least for spontaneous speech. The study's results support early introduction of PPS and use of LT and LTA for monitoring and evaluating it, first using LT for raising learners' awareness of their own production. Then, as learners become familiar with rules for PPS prediction, instructors may move them on to the prediction and self-correction stages

(annotation). Intonation is closely related to PPS, thus these two topics should be covered together.

FW and Link seemed to improve with multiple rehearsals for each of the three main strategy types, possibly due to increased familiarity with the content. Thus rehearsal may be an important strategy to incorporate into instruction on these features.

The learners in this study had already achieved a high level of accuracy for vowel reduction in content words (92%) and word stress (95%). However, this level of accuracy may be characteristic of high-frequency words in the learners' lexicons and may not represent their overall competence. Instructors should assess learners' ability to reduce unstressed vowels and produce accurate word stress in both familiar and unfamiliar (and relevant) content words and prioritize this target accordingly.

How to teach the strategies. Before instruction begins, language teachers must determine how strategy instruction will be integrated into the class, including amount of time spent on it, how tasks will be designed to incorporate strategy use, how to assess learners' effectiveness using the strategies, and how to intervene for the less successful learners.

During the first week, a diagnostic test should be administered to determine the targets for which learners will derive the greatest benefits from self-monitoring.

If pronunciation strategy instruction will be a significant focus, have students begin using the target strategies early in the term and evaluate their performance regularly. This helps determine which learners can move ahead more rapidly with less guidance and which ones need more help using the strategies appropriately. As described in Chapter 4 and illustrated in Appendix A, instruction for the strategies and pronunciation features are closely related. Classroom and homework tasks may be designed to reinforce both topics.

Lower proficiency learners will need systematic guidance when using annotation. Assignments should have a clear structure with step by step instructions and teacher feedback built in regularly to ensure that learners are making correct predictions.

How can we match strategy use to learners? Though this study was not designed to explore why strategy use was more effective for some learners and not others, the results clearly indicate that this was the case. As language teachers, we want to provide instruction that meets our learners' needs, learning styles and preferences, and proficiency level, to the extent possible. Chamot and Kupper (1989) describe a very practical approach to identifying students' strategy preferences and current ability, including asking learners to describe their strategy use, through small-group discussions in class where peers help each other describe their learning approaches, and through the instructor's observations of learners' strengths and weaknesses in using the target strategies. Observing whether learners tend to prefer more intuitive or analytic approaches to learning or whether they are fluency focused or accuracy focused may guide material and strategy selection.

The L, LT, and LTA strategy types were described in this study as a type of covert rehearsal. However, these tasks can be made into communicative activities, in which learners listen to each other's recordings and prepare transcriptions together. Learners also may create together a recording of a communicative task and then listen, transcribe, annotate, and revise the recording until they have achieved targetlike pronunciation (Lynch, 2007).

Individualizing instruction. Language instructors typically find a range of proficiency levels within any given class, even when placement tests and progress through a curriculum determine who enrolls in a course. Evaluating entering proficiency can be a way to target instruction to match individuals' pronunciation needs, learning styles, and strategy use, and

possibly offer principles for assigning grades. The tools for assessing language proficiency and strategy use were noted previously: Gather information from diagnostic tests during the first week of instruction, have students individually and with peers identify their own strategies and approaches for language learning, and observe students using the strategies.

Setting a foundation early regarding the target pronunciation features and strategies can make individualizing instruction easier and more accurate. If learners understand definitions of PPS, MU boundaries, and other targets, and if they know why and how to use the strategies, then advanced learners can be pushed ahead to more difficult targets and learners who make slower progress can be scaffolded in their work with the targets and strategies.

The target strategies follow a progression, as outlined in Figure 1 (Chapter 2). Instructors can vary which types of self-monitoring learners use. For example, for features that appear to be easier, such as MU boundaries, learners should be able to use the full strategy complement: LTA. But if a learner is having trouble perceiving his or her production or is having difficulty applying rules, then the annotation stage should be delayed. Learners who master a pronunciation target can move on to more difficult ones and may transition to annotation only when they are ready. And as learners approach mastery, they may find that listening only is sufficient for monitoring and correcting their production. By helping learners focus on their most pressing pronunciation concerns and helping them use the strategies that are appropriate for their level, instructors can effectively individualize instruction.

Using transcription to understand what learners are doing. As a language instructor, I found that listening to and transcribing the students' speech data informed my teaching in meaningful ways. I learned a great deal about occurrence of the nontargetlike features of their speech, such as which specific function words they were not reducing, in which environments

segmental errors tended to occur, and the extent to which lack of segmental versus suprasegmental accuracy seemed to contribute more significantly to intelligibility. For example, P3 (Mandarin L1) and P8 (Korean L1) each tended to use very few function words compared to the other participants. This seemed to have an impact on the amount of linking that was possible. Additionally, I discovered that learners typically recycled the same vocabulary and used few pronouns, thus decreasing the coherence of their discourse.

Of course language instructors typically do not have time to transcribe large portions of their students' oral production. However, transcription certainly could be useful for teachers less experienced in teaching pronunciation, for those working with a new group of students from a less familiar L1 background, or in a situation in which identifying specific learner needs is useful for addressing immediate intelligibility issues, as when a teacher is helping a student prepare for a conference presentation.

In the next section, I discuss how a different type of proficiency score can be used to identify learners' proficiency and anticipate the amount and type of progress learners may make in a semester.

Using improvement scores to evaluate learner accuracy gains. In Chapter 6, I used difference scores (described in Chapter 5) when discussing the results of this study. These scores represent absolute differences between each rehearsal and the baseline and are appropriate for evaluating total group results in a repeated measures ANOVA design. However, absolute scores do not provide insight into how many of the nontarget tokens were corrected following self-monitoring, nor do such scores indicate how close the learner is to achieving 100% accuracy. For example, does a 10% gain mean the learner started at 90% baseline accuracy and subsequently corrected all nontarget features? Or perhaps she started at a 50% baseline and reached 60%

accuracy at R3? Is it appropriate to expect learners at the high and low ends of the proficiency range to make similar gains?

W. B. Dickerson (1997), M. K. Hahn (2002), and Sardegna (2009) found that larger accuracy increases (based on difference scores) tended to occur for those participants whose entering proficiency was highest, most likely because they had developed greater control over the target pronunciation features and thus could make larger gains. Similarly, participants starting at a lower baseline percentage tended to make smaller increases, due to lower entering proficiency and limited mastery of the targets. As noted in Chapter 6, results of the current study did not follow this pattern. Instead, learners with lower starting proficiencies made larger gains. But the pattern itself is not as important to this discussion, rather the fact that high and low proficiency learners perform differently is of interest here.

Here is an illustration of how the same difference score may be interpreted differently by looking at the number of nontarget features corrected. For example, P2's overall difference score was 7.8%, meaning that following self-monitoring, his accuracy increased from a baseline of 66% to 73.8%. P6 had the same difference score (7.8%), but this moved her from a baseline of 60.5% to 68.3%. Thus both learners increased their accuracy the same amount, but ended up at different overall accuracy levels:

Participant	Difference score	Baseline accuracy
P2	7.8%	66.0%
P6	7.8%	60.5%

Another way to present accuracy gains or losses while integrating differences in proficiency level is to look at relative improvement: the percentage of nontarget tokens that were corrected following self-monitoring. Let us look more closely at the data for P2 and P6:

For P2, the percentage of nontarget tokens at ML3 was 34%:

100	%	(maximum accuracy)
- 66	%	(baseline accuracy)
34	%	(percentage of tokens that are nontargetlike)

P2 corrected 7.8% of the nontargetlike tokens, or 23%:

7.8% / 34% = 23%

For P6, a difference score of 7.8% represents 19.8% improvement (7.8%/39.5%).

P15 and P4 also had similar difference scores (7.7% and 7.6%, respectively; Figure 9). A comparison of P2, P4, P6, and P15 shows that these nearly identical difference scores represent a range of improvement from 17.6% to 24.4%. The goal for learners is to achieve 100% accuracy. Figure 9 shows how close participants came to doing so. Remember that participant numbers refer to their baseline accuracy.

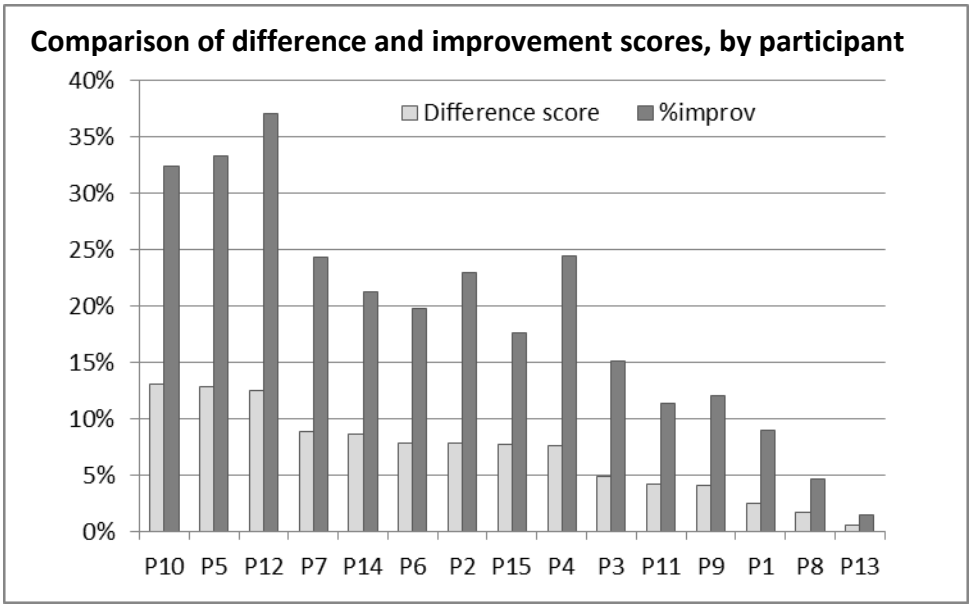


Figure 9. Comparison of difference scores and improvement scores for each learner from ML3 to R3, for all targets and strategies. Difference scores represent absolute gains or losses, but improvement scores are relative. Participants with similar difference scores may have substantially different improvement scores if their proficiency levels also are different.

Following is a summary of the advantages and disadvantages of each type of score. For research purposes, difference scores are preferred, and for pedagogical purposes, improvement scores offer several benefits.

Difference scores = Post-instruction score – pre-instruction score	Improvement scores = difference score / (1 - baseline accuracy)
Advantages: Same metric used in most comparable studies, making it easier to compare results. Appropriate for use in computing inferential statistics.	Advantages: Places learners at the same starting point: how close the learner is to 100% accuracy. Pedagogical value: Acknowledges differences in amount of progress between high and low groups, thus helping instructors and learners have realistic expectations for pronunciation gains. Can be used to develop grading criteria that take into account expectations for reasonable accuracy gains.
Disadvantages: Magnitude of differences in performance are not apparent between learners with the same difference scores but different starting proficiencies.	Disadvantages: Rarely used as a metric in other studies. Differences in performance between learners at different starting proficiencies are not apparent. Scores for learners at the extremes are skewed. Learners with baselines near 100% will have especially large % improvement scores given the difference score size. 100% improvement could represent the correction of only 2 or 3 tokens. A learner with a much lower baseline score (e.g., 50%) might correct 2 or 3 tokens and have an improvement score of below 10%.

Next I illustrate with the current study's data how improvement scores provide another approach to interpreting the relationship between proficiency and learner performance. Following the rationale and procedures outlined in W. B. Dickerson (1997), I divided the participants into Low and High proficiency groups in the following way: I ranked participants according to their overall baseline accuracy scores from ML1. I identified a gap between the seventh and eighth participants, resulting in seven High and eight Low participants (Table 31).

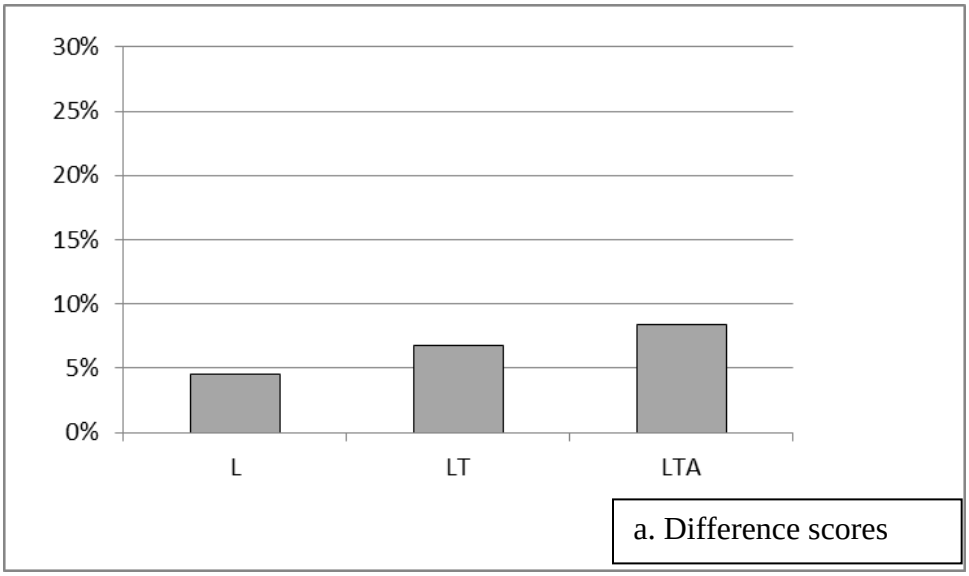
Table 31

High and Low Proficiency Groups, Based on Pre-Instruction Baseline

Group	Participant	ML1 baseline accuracy
High	P1	73.3%
	P2	72.4%
	P3	70.9%
	P4	69.3%
	P5	68.6%
	P6	67.9%
	P7	66.9%
Low	P8	64.5%
	P9	64.2%
	P10	62.9%
	P11	61.5%
	P12	60.7%
	P13	60.0%
	P14	58.7%
	P15	52.9%

When proficiency is taken into account and percentage improvement is evaluated, a slight distinction emerges between the High and Low groups (Figure 10). Figure 10a shows overall group difference scores for the three strategy types. In Figure 10b, the same data are broken down by proficiency group using improvement scores. Both groups make accuracy improvements with each strategy type. However, we can now see more clearly that the High

group benefits most from LTA, whereas for the Low group, LT is the most effective strategy type. As mentioned in previous sections, lower proficiency learners may be less ready to use annotation effectively: They may make incorrect predictions due to less mastery of pronunciation rules. Higher learners may benefit from annotation: It may help them correct the most difficult nontarget features.



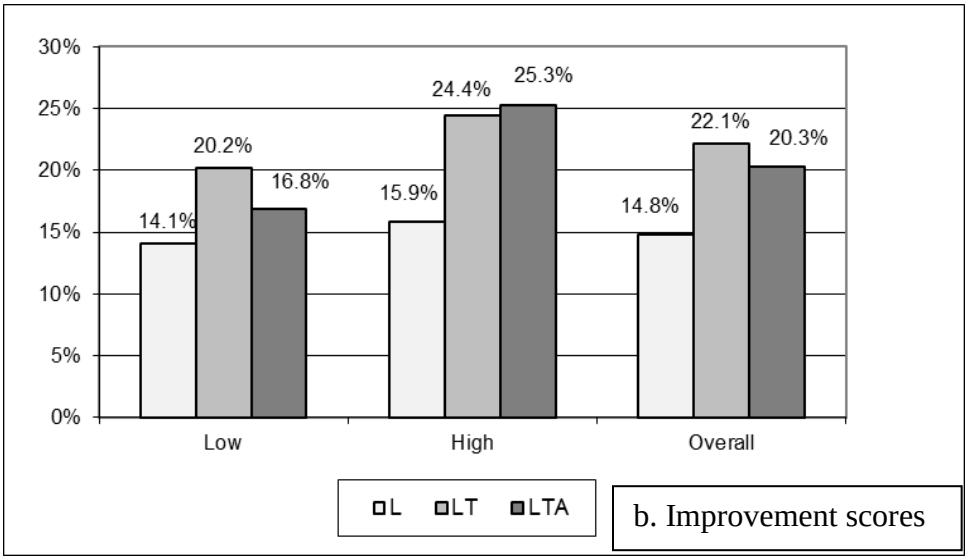


Figure 10. (a) Grand means for difference scores for each strategy type. Data are combined for all participants, targets, and rehearsals. A significant difference exists between L and LTA (p = .05), and LTA is the most effective strategy type. (b) Overall percentage improvement for all participants and targets in comparison to High and Low proficiency groups. Both groups benefit from all three strategies. LTA is most effective for High, and LT is the most effective strategy type for Low.

Looking at individual improvement scores by strategy type helps us see in greater detail how learners' proficiency related to strategy use, and the picture is adjusted slightly. High learners tended to make strong gains across all strategy types and low learners tended to do their

best with only one or two types Table 32. And comparing improvement scores in Table 32 to the differences scores in Table 33, we see that for P14, LT is actually the more effective strategy. And P15's and P3's difference scores are nearly equal (10.7% and 10.1%, respectively), but P15's improvement score (progress toward 100% accuracy) is greater than P3's (32.1% vs. 20.7%, respectively). These two learners also fell into different proficiency groups.

Table 32

Participants' Improvement Scores, by Strategy Type and Proficiency Group

Participant	Proficiency group	L	LT	LTA
P5	High	**42.1%**	30.9%	27.3%
P7	High	**29.4%**	20.7%	21.6%
P11	Low	**20.0%**	7.8%	4.7%
P1	High	**18.2%**	4.5%	8.9%
P10	Low	26.6%	**41.4%**	29.8%
P4	High	-8.3%	**37.5%**	35.9%
P2	High	12.7%	**34.0%**	23.7%
P6	High	1.7%	**30.7%**	23.2%
P9	Low	-4.5%	**24.6%**	9.7%
P14	Low	13.5%	**24.6%**	23.5%
P13	Low	-6.7%	**17.4%**	-6.4%
P8	Low	0.0%	**13.2%**	3.4%
P12	Low	40.0%	17.1%	**45.9%**
P3	High	8.9%	3.7%	**32.1%**
P15	Low	17.5%	13.9%	**20.7%**

Note. Each participant's greatest improvement score appears in boldface.

Table 33

Participants' Difference Scores, by Strategy Type and Proficiency Group

Participant	L	LT	LTA
P5	**15.4%**	12.7%	10.4%
P7	**12.7%**	7.4%	6.5%
P11	**8.0%**	2.8%	1.6%
P1	**3.3%**	1.5%	2.7%
P10	11.6%	**16.6%**	11.2%
P6	0.6%	**13.3%**	9.4%
P4	-2.3%	**12.2%**	11.4%
P2	4.5%	**10.0%**	9.0%
P9	-1.3%	**9.3%**	3.3%
P13	-2.3%	**7.3%**	-2.6%
P8	0.0%	**4.0%**	1.4%
P12	13.3%	4.5%	**18.6%**
P14	4.5%	9.8%	**11.2%**
P15	7.2%	5.3%	**10.7%**
P3	2.5%	1.4%	**10.1%**

Note. Each participant's greatest difference score appears in boldface.

An important implication of using improvement scores is that they can help instructors create more realistic and appropriate expectations for learner improvement over the course of a semester. As noted by W. B. Dickerson (1997), if a sufficient spread of proficiency scores exists, end-of-semester grading scales may be adjusted to recognize the fact that lower students' gains may be different from those achieved by higher students. The bottom line is that factors beyond the learners' control often affect the amount of progress a learner makes in a semester (see Chapters 1 and 2): language and auditory aptitude, working memory limits, and developmental readiness. But as language teachers, we do not want to penalize (with lower grades) a group whose gains are smaller when in fact the gain size may have been exactly what was reasonable for someone at a lower proficiency level to accomplish.

Future research

The outcomes of this study and others preceding it tell us that strategy-based instruction is effective for many, if not most, learners, and is worthwhile for a range of L2 skills and particularly for increasing pronunciation accuracy. Several sets of questions need further investigation to fill in the many remaining gaps in what we know about the effectiveness of helping language learners develop strategic competence. Answers to these questions will help language instructors and learners prioritize pronunciation improvement efforts.

1. How does use of self-monitoring strategies relate to accuracy in spontaneous production? How do we know that strategy instruction leads to greater accuracy in the long term? The ultimate outcome that language instructors strive for is that students be able to produce pronunciation targets accurately and spontaneously, without conscious self-monitoring. Longitudinal work that links specific strategies to specific pronunciation targets and measures prolonged strategy use is one research course, similar to the work done by M. K. Hahn (2002) and Sardegna (2009). A modified research design is needed to give us a better sense of how self-monitoring works. Perhaps a focus on one target feature at a time would give a better understanding of strategy effectiveness. Another interesting way to look at this could be to assess learners' progress more frequently and on a variety of task

types (ranging from controlled reading tasks to spontaneous responses to prompts) during the semester in order to identify changes. Both classroom-based and experimental studies would be useful for observing how learners use pronunciation strategies to make improvements on each specific target.

2. A second key question is to what extent is improvement for a given pronunciation target linked to increased intelligibility? Previous researchers have found that prosodic features such as stress, rhythm, intonation, and phrasing may contribute greatly to intelligibility (Anderson-Hsieh & Koehler, 1988; Anderson-Hsieh et al., 1992; Johansson, 1978; Pickering, 2001; Tajima, Port, & Dalby, 1997; Wennerstrom, 1998). Though some studies have isolated effects of a single pronunciation feature (intonation by Wennerstrom and Pickering; PPS by L. D. Hahn (1999, 2004); segmentals and word stress by Isaacs, 2006), conclusive information about the relative importance of various features is not definitive. And though one clear answer may not be possible, knowing more about relative contributions of each pronunciation feature to intelligibility would be very useful. An additional step in the current study could provide useful information toward that goal, by having speech samples rated holistically for comprehensibility. These ratings could be compared to accuracy scores for the various suprasegmental features to highlight which targets and in which combinations were most critical for intelligibility.

3. Additional data are needed to better evaluate strategy effectiveness for improving word stress, construction stress, and vowel reduction in content words. To obtain appropriate data, tasks involving reading, describing pictures, or other carefully controlled prompts are needed to elicit sufficient and varied tokens for each target.

4. As noted earlier, additional research is needed on the effectiveness of strategy use by learners from other L1 backgrounds and lower L2 proficiency levels.

5. Another interesting question is, can these strategies be used effectively for segmental features? My suspicion is that more interactive strategies might be more effective, given that learners often have difficulty perceiving what they are doing, particularly with the more challenging sound contrasts, such as l/r or iy/ɪ. For example, I do not recall any of the participants in this study noting nontarget segments, such as writing down that they said “sink” when they meant to say “think.” English orthography is an important factor in learners’ segmental pronunciation, thus finding out whether or not transcription is effective would be helpful.

6. Data from the current study merit further examination and may offer some useful insights into L2 learner behavior. For example, (a) how accurate were learners’ transcriptions in terms of reflecting what the learner produced in the minilecture; (b) how accurately learners corrected their transcripts (LTA); and (c) how well the rehearsals matched what learners marked on their transcriptions?

Conclusion

The results of this study move us a few steps further in our understanding of the merits of strategy instruction for L2 pronunciation improvement. Research on this topic is limited, and most research has focused on identifying LLS taxonomies. Pronunciation strategy research indicates that learners can use self-monitoring strategies for improving pronunciation accuracy, but most studies have focused on holistic/global pronunciation improvement, rather than on identifying how the target strategies affect accuracy on specific pronunciation features. This study helps fill some of the gaps in our understanding of the utility of strategy-based instruction and the extent to which specific pronunciation strategies are useful for correcting an even wider range of suprasegmentals than previously studied. No other studies have reported token counts to demonstrate accuracy change for message unit boundaries, linking, reduction of unstressed function words, and intonation. Thus the current study provides greater insight to the detailed pronunciation behavior of L2 learners when speaking spontaneously.

Knowing that specific self-monitoring strategy combinations may be more suitable for specific pronunciation targets can aid language teachers as they target their instruction. Self-monitoring skills may be useful for learners interested in enhancing their study practices for traditional and online classes and for post-instruction study.

Self-monitoring strategies are not the only strategies that learners should use. Nor is explicit attention to form sufficient for learners to improve pronunciation. Rather the primary value of the current study is the finding that, for many learners, the ability to effectively self-monitor develops relatively quickly following a period of systematic training. The resulting increase in sensitivity to the important L2 features in one's own speech may, over time, result in improved L2 production and less reliance on conscious use of meta-cognitive learning strategies.

References

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly, 18*(1), 71-85.

Akita, M. (2005). The effectiveness of a prosody-oriented approach in L2 perception and production. In A. Brugos, M. R. Clark-Cotton, & S. Ha (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development: Vol. 1* (pp. 24-36). Somerville, MA: Cascadilla Press.

Anani, M. (1989). Incorrect stress placement in the case of Arab learners of English. *International Review of Applied Linguistics, 27*, 15-21.

Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The Modern Language Journal, 75*(4), 460-472.

Anderson, P. (1993). *The interstress interval as an indicator of perceived intelligibility among nonnative speakers of English* (Unpublished doctoral dissertation). Retrieved from ProQuest Dissertation Abstracts. (AAC 9325699)

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

Anderson-Hsieh, J. (1992). Using electronic visual feedback to teach suprasegmentals. *System, 20*(1), 51–62.

Anderson-Hsieh, J. (1994). Interpreting visual feedback on suprasegmentals in computer assisted pronunciation instruction. *CALICO Journal, 11*(4), 5-22.

Anderson-Hsieh, J., & Koehler, K. (1988). The effects of speaking rate and foreign accent on native speaker comprehension. *Language Learning, 38*(4), 561-613.

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.

Archibald, J. (1995). A longitudinal study of the acquisition of English stress. *Calgary Working Papers in Linguistics, 17*, 1-10.

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

Atkinson, J. M, & Heritage, J. (Eds). (1984). *Structures of social action: Studies in conversation analysis*. Cambridge, UK: Cambridge University Press.

Aufderhaar, C. (2004). The influence of using discourse analysis techniques on the filtered speech of authentic audio text to improve pronunciation (Unpublished doctoral dissertation). University of Cincinnati, OH.

Avery, P., & Ehrlich, S. (1987). Specific pronunciation problems. *TESL Talk, 17*, 82-115.

Aziz, Y. Y. (1980). Some problems of English word-stress for the Iraqi learner. *English Language Teaching Journal, 34*, 104-109.

Baptista, B. O. (1989). Strategies for the prediction of English word stress. *International Review of Applied Linguistics, 27*, 1-14.

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

Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America, 114*(3), 1600-1610.

Blanche, P. (2004). Using dictation to teach pronunciation. *Linguagem & Ensino, 7*(1), 175-191.

Bolinger, D. L. (1986). *Intonation and its parts: Melody in spoken English.* Stanford, CA: Stanford University Press.

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

Breiner-Sanders, K. E., Lowe, P., Miles, J., & Swender, E. (1999). ACTFL proficiency guidelines—Speaking: Revised 1999. *Foreign Language Annals, 33*(1), 13-18.

Breun, J. (2001a). Strategies for success: Profiling the effective learner of German. *Foreign Language Annals, 34*(3), 216-225.

Breun, J. (2001b). The parallel development of oral proficiency and use of language learning strategies. *Teaching German, (34)*2, 158-168.

Brown, A. (1988). Functional load and the teaching of pronunciation. *TESOL Quarterly, 22*(4), 593-606.

Brown, H. D. (2007). *Teaching by principles: An interactive approach to language pedagogy* (3rd ed). Englewood Cliffs, NJ: Prentice Hall.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23-48). Harlow: Longman.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge, UK: Cambridge University Press.

Chamot, A. U., & Kupper, L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals, 22*(1), 13-24.

Chang, J. (2001). Chinese speakers. In M. Swan & B. Smith (Eds.), *Learner English* (2nd ed.), (pp. 310-324). Cambridge, UK: Cambridge University Press.

Chang, M. G. (2006). *Pronunciation instruction, learner awareness and development*. (Unpublished doctoral dissertation). University of Toronto, Canada.

Chang, S. J. (1990). *A study of language learning behaviors of Chinese students at the University of Georgia and the relation of those behaviors to oral proficiency and other factors* (Unpublished doctoral dissertation). University of Georgia, Athens.

Chela-Flores, B. (2001) Pronunciation and language learning: An integrative approach. *International Review of Applied Linguistics, 39*(2), 85-111.

Chen, Y. (2001). A study of sentence stress production in Mandarin speakers of American English. *Journal of the Acoustical Society of America, 109*, 1681-1690.

Chen, Y. M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research, 12*(2), 235-262.

Chesterfield, R., & Chesterfield, K. B. (1985). Natural order in children's use of L2 learning strategies. *Applied Linguistics, 6*(1), 45-59.

Clark, J., & Yallop, C. (1995). *An introduction to phonetics and phonology* (2nd ed.). Malden, MA: Blackwell.

Clennell, C. (1999). Promoting pragmatic awareness and spoken discourse skills with EAP classes. *ELT Journal, 53*(2), 83-91.

Cohen, A. D., Weaver, S., & Li, T. (1995). *The impact of strategies-based instruction on speaking a foreign language*. Minneapolis, MN: University of Minnesota.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and nonnative listeners. *Language and Speech, 45*(2), 207-228.

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

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

Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge, UK: Cambridge University Press.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141-201.

Cutler, A., Weber, A., & Otake, T. (2006). Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics, 34*(2), 269-284.

deBot, K. (1983). Visual feedback of intonation I: Effectiveness and induced practice behavior. *Language and Speech, 26*(4), 331-350.

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

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39*(3), 379-397.

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

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

Dickerson, L. J. (1975). The learner's interlanguage as a system of variable rules. *TESOL Quarterly, 9*, 401-407.

Dickerson, L. J., & Dickerson, W. B. (1982). *Learning English pronunciation: Vowels*. Unpublished manuscript.

Dickerson, L. J., & Dickerson, W. B. (1983). *Learning English pronunciation: Consonants.* Unpublished manuscript.

Dickerson, W. B. (1978). English orthography: A guide to word stress and vowel quality. *International Review of Applied Linguistics, 16*, 127-147.

Dickerson, W. B. (1984). The role of formal rules in pronunciation instruction. In J. Handscombe, R. Orem, & B. Taylor (Eds.), *On TESOL '83* (pp. 135-148). Washington, D.C.: TESOL.

Dickerson, W. B. (1987). Explicit rules and developing interlanguage phonology. In A. James & J. Leather (Eds.), *Sound patterns in second language acquisition* (pp. 121-140). Berlin: Foris.

Dickerson, W. B. (1989). *Stress in the speech stream: The rhythm of spoken English* [CD-ROM]. Urbana, IL: Author.

Dickerson, W. B. (1994). Empowering students with predictive skills. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New views, new directions* (pp. 19-35). Arlington, VA: TESOL.

Dickerson, W. B. (1997). Grading pronunciation progress: The roles of absolute and relative scales. In W. Davis, J. Smith, & R. Smith, *Ready to teach: Graduate teaching assistants prepare for today and for tomorrow* (pp. 159-166). Stillwater, OK: New Forums Press.

Dickerson, W. B. (2000). *Covert rehearsal as a bridge to accurate fluency.* Presented at the 34th Annual Meeting of the Teachers of English to Speakers of Other Languages, Vancouver, BC.

Eckstein, G. T. (2007). *A correlation of pronunciation learning strategies with spontaneous English pronunciation of adult ESL learners* (Unpublished master's thesis). Brigham Young University, Provo, UT.

Esling, J. H., & Wong, R. F. (1983). Voice quality settings and the teaching of pronunciation. *TESOL Quarterly, 17*(1), 89-94.

Ewing, K. S. (2002). *The effects of oral reading on the intonation and past tense verb use of adult non-native speakers of English* (Unpublished doctoral dissertation). University of Washington, Seattle.

Fagundes, R. B. L. F. (1994). *An investigation of selected variables as predictors of success on the SPEAK and TEACH tests* (Unpublished doctoral dissertation). Iowa State University, Ames.

Feyten, C. M., Flaitz, J. J., & LaRocca, M. A. (1999). Consciousness raising and strategy use. *Applied Language Learning, 10*(1 & 2), 15-38.

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

Firth, S. (1992). Developing self-correcting and self-monitoring strategies. In P. Avery & S. Ehrlich (Eds.), *Teaching American English pronunciation* (pp. 215-219). Oxford: Oxford University Press.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18,* 299-323.

Fudge, E. (1984). *English word-stress*. London, UK: George Allen & Unwin.

Gilbert, J.B. (1993) *Clear speech: Pronunciation and listening comprehension in North American English.* New York, NY: Cambridge University Press.

Gorsuch, G. J. (2001). Testing textbook theories and tests: The case of suprasegmentals in a pronunciation textbook. *System, 29*, 119-136.

Grenfell, M., & Macaro, E. (2007). Claims and critiques. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. vii-viii). Oxford: Oxford University Press.

Gu, P. Y. (2007). Foreword. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. vii-viii). Oxford: Oxford University Press.

Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning, 46*(4), 643-679.

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

Hahn, L. D. (1994). The stress of compound nouns: Linguistic considerations and pedagogical considerations. *IDEAL, 7*, 67-78.

Hahn, L. D. (1999). *Native speakers' reactions to non-native speech in English discourse* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.

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

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

Hahn, M. K. (2002). *The persistence of learned primary phrase stress patterns among learners of English* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.

Harris, D. F. (2003). The effects of training on the pronunciation of Mandarin Chinese speakers. *Dissertation Abstracts International: Section B. Sciences and Engineering, 63*(8), 3687.

Haslam, N. (2010). *The relationship of three L2 learning factors with pronunciation proficiency: Language aptitude, strategy use, and learning context* (Unpublished master's thesis). Brigham Young University, Provo, UT.

Hassan, X., Macaro, E., Mason, D., Nye, G., Smith, P., & Vanderplank, R. (2005). Strategy training in language learning: A systematic review of available research. In *Research Evidence in Education Library*. London, UK: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Hsiao, T. Y., & Oxford, R. L. (2002). Comparing theories of language learning strategies: A confirmatory factor analysis. *The Modern Language Journal*, *86*(3), 368-383.

Hua, T. (2003). *The acquisition of English speech rhythm by adult Chinese ESL and EFL learners* (Unpublished doctoral dissertation). University of Hawaii at Manoa.

Isaacs, T. (2006). *Towards defining a valid assessment criterion of pronunciation proficiency in non-native English speaking graduate students* (Unpublished master's thesis). McGill University, Montreal, Canada.

Izumi, S. (2003). Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics, 24*(2), 168-196.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics, 23*(1), 83-103.

Jilka, M. (2000). *The contribution of intonation to the perception of foreign accent* (Unpublished doctoral dissertation). University of Stuttgart, Germany.

Johansson, S. (1978). Studies of error gravity: Native reactions to errors produced by Swedish learners of English. *Gothenburg Studies in English, 44*, 1-138.

Juffs, A. (1990). Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *International Review of Applied Linguistics, 28*, 99-117.

Juffs, A., & Rodriquez, G. (2007). *Working memory in context: Differential effects on comprehension of relative clauses and anaphor binding.* Presented at the 30th annual meeting of the Second Language Research Forum, Urbana, IL.

Kawagoe, I. (2003). Acquisition of English word stress by Japanese learners. In J. M. Liceras, H. Zobl, & H. Goodluck (Eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference* (pp. 161-167). Somerville, MA: Cascadilla.

Kazdin, A. E. (1982). *Single-case research design: Methods for clinical and applied settings.* New York, NY: Oxford University Press.

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

Kenworthy, J. (1987). *Teaching English pronunciation.* London, UK: Longman.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed). Pacific Grove, CA: Brooks/Cole.

Ladd, D. R. (1996). *Intonational phonology.* Cambridge, UK: Cambridge University Press.

Lai, Y. C. (2009). Language learning strategy use and English proficiency of university freshmen in Taiwan. *TESOL Quarterly, 43*(2), 255-280.

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3), 369-377.

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

Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford: Oxford University Press.

Lynch, T. (2007). Learning from the transcripts of an oral communication task. *ELT Journal, 61*(4), 311-320.

Lynch, T., & Maclean, J. (2001). 'A case of exercising': Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 141-162). Harlow: Longman.

Macaro, E. (2006). Strategies for language learning and for language use: Revising the theoretical framework. *The Modern Language Journal, 90*(3), 320-337.

Macdonald, D. M., Yule, G., & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning, 44*(1), 75-100.

MacKay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America, 110*, 516-528.

McDonough, S. H. (1995). *Strategy and skill in learning a foreign language.* London, UK: Edward Arnold.

Mennim, P. (2003). Rehearsed oral L2 output and reactive focus on form. *ELT Journal, 57*(2), 130-138.

Mennim, P. (2007). Long-term effects of noticing on oral output. *Language Teaching Research, 11*(3), 265-280.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 101*(2), 343-352. Retrieved from http://www.psych.utoronto.ca/users/peterson/psy430s2001/Miller%20GA%20Magical%20Seven%20Psych%20Review%201955.pdf

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481-520.

Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*, 81–108.

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*(1), 111-131.

Murphy, J. (2004). Attending to word-stress while learning new vocabulary. *English for Specific Purposes, 23*, 67-83.

Myers, S. A. (1995). Using written text to teach oral skills: An ITA training class using field-specific materials. *English for Specific Purposes, 14*, 231-245.

Naiman, N., Frohlich, M., Stern, H. H., & Todesco, A. (1996). *The good language learner.* Clevedon, UK: Multilingual Matters.

Nguyen, T. T. A., & Ingram, J. (2005). Vietnamese acquisition of English word stress. *TESOL Quarterly, 39*, 309-319.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition.* Cambridge: Cambridge University Press.

O'Malley, J. M., Chamot, A. U., Stewner-Manzanares, G., Kupper, L., & Russo, R. P. (1985). Learning strategies used by beginning and intermediate ESL students. *Language Learning, 35*(1), 21-46.

Osburne, A. G. (2003). Pronunciation strategies of advanced ESOL learners. *International Review of Applied Linguistics, 41*(2), 131.

Oxford, R. L. (1989). Language learning strategies, the communicative approach, and their classroom implications. *Foreign Language Annals, 22*(1), 291-300.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know.* New York, NY: Newbury House.

Oxford, R., & Nyikos, M. (1989). Variables affecting choice of language learning strategies by university students. *The Modern Language Journal, 73*(3), 291-300.

Pawlak, M. (2010). Designing and piloting a tool for the measurement of the use of pronunciation learning strategies. *Research in Language, 8,* 189-202.

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

Perez-Gamboa, J. (1989). Spanish-speaking students' common English mistakes in stress and intonation. *Lenguaje y Ciencias, 29*, 11-16.

Peterson, S. S. (1997). *Pronunciation learning strategies and learning strategies related to pronunciation ability in American university students studying Spanish* (Unpublished doctoral dissertation). The Ohio State University, Columbus.

Peterson, S. S. (2000). *Pronunciation learning strategies: A first look.* Unpublished research report. (ERIC Document Reproduction Service ED 450 599; FL 026 618). Retrieved from http://www.eric.ed.gov/PDFS/ED450599.pdf

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

Purpura, J. E. (1997). An analysis of the relationships between test takers' cognitive and metacognitive strategy use and second language test performance. *Language Learning, 47*(2), 289-325.

Ramirez Verdugo, D. (2006). A study of intonation awareness and learning in non-native speakers of English. *Language Awareness, 2006, 15*(3), 141-159.

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

Ridley, J. (1997). *Reflection and strategies in foreign language learning.* Frankfurt: Peter Lang.

Riney, T. J., & Flege, J. E. (1998). Changes over time in global foreign accent and liquid identifiability and accuracy. *Studies in Second Language Acquisition, 20*, 213–243.

Rubin, J. (1975). What the "good language learner" can teach us. *TESOL Quarterly, 9*(1), 41-51.

Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics, 11*(2), 118-131.

Sardegna, V. G. (2009). *Improving English stress through pronunciation learning strategies* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign.

Scarcella, R. C., & Oxford, R. L. (1994). Second language pronunciation: State of the art in instruction. *System, 22*(2), 221-230.

Seferolug, G. (2005). Improving students' pronunciation through accent reduction software. *British Journal of Educational Technology, 36*(2), 303–316.

Shattuck-Hufnagel, S., & Turk, A. (1996) A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research, 25*(2), 193-247.

Slowiaczek, L. M. (1990). Effects of lexical stress in auditory word recognition. *Language and Speech, 33*(1), 47-68.

Smith, J., & Beckman, B. (2010). Noticing-reformulation task as a stimulus towards continuous autonomous phonological development. *New Zealand Studies in Applied Linguistics, 16*(1), 35-50.

Swan, M., & Smith, B. (Eds.). (2001). *Learner English* (2nd ed.). Cambridge, UK: Cambridge University Press.

Stern, H. H. (1975). What can we learn from the good language learner? *The Canadian Modern Language Review, 31*(4), 304-318.

Takeuchi, O., Griffiths, C., & Coyle, D. (2007). Applying strategies to contexts: The role of individual, situational, and group differences. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. 69-92). Oxford: Oxford University Press.

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

Tarone, E. (1985). Variability in interlanguage use: A study of style shifting in morphology and syntax. *Language Learning, 35,* 373-403.

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

Vann, R. J., & Abraham, R. G. (1990). Strategies of unsuccessful language learners. *TESOL Quarterly, 24*(2), 177-198.

van Santen, J. P. H., & Shih, C. (2000). Suprasegmental and segmental timing models in Mandarin Chinese and American English. *Journal of the Acoustical Society of America, 107,* 1012-1026.

Vitanova, G., & Miller, A. (2002). Reflective practice in pronunciation learning. *The Internet TESL Journal*. Retrieved from http://iteslj.org/Articles/Vitanova-Pronunciation.html

Wennerstrom, A. (1998). Intonation as cohesion in academic discourse. *Studies in Second Language Acquisition, 20*(1), 1-25.

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

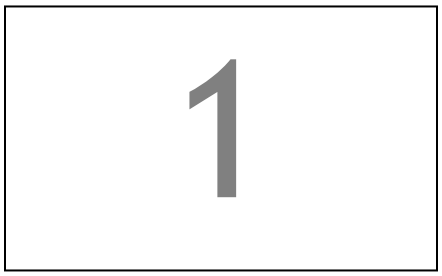
Willingham, D. B., & Goedert-Eschmann, K. (1999). The relation between implicit and explicit learning: Evidence for parallel development. *Psychological Science*, (*10)*6, 531-534.

Appendix A

Task Instructions for Participants

ML3 Self-Evaluation Instructions

Part 1



Instructions

First, open Audacity and size its window to fit the bottom half of your desktop.

Place your CD in your computer.

Go to "My Computer", open the CD folder, and then open the .mp3 file called "your name_Part_1". Pause the player.

Organize the player and Audacity on your desktop so you can see both at the same time.

When you are ready, click "start" for your ML1 recording and follow the instructions.

See the reverse side of this sheet for a printed version of the instructions you will hear on the CD.

Part A

Hi, this is Sue. Today you will listen to part of your ML3 recording.

I will give you instructions as we go along. I will tell you when it's time to record your voice. First, I will ask you to listen several times to part of your minilecture.

1. First please listen to the following part of your lecture, to familiarize yourself with what you said.

2. This time please listen again to the same recording. I want you to focus on how you used message unit breaks and where you paused and think about how you could pause differently.

3. This time please listen again to the same recording. I want you to focus on how you used primary phrase stress and intonation. Please think about how you could improve these features as you're listening.

4. This time please listen again to the same recording. I want you to focus on how you used word stress and compound noun or compound number stress, if you used those two. Please think about how you could improve these features.

5. This time listen again to the same recording. I want you to focus on your rhythm, including vowel reduction and using clear stress in the right location. Please think about how you could improve these features.

Please click the red record button in Audacity. Do not press the yellow stop button until I tell you to do so.

Next, you will listen to the recording in shorter chunks. After listening to a chunk, repeat, out loud, what you said originally. After speaking, think about your pronunciation accuracy. I'll ask you to listen and repeat a total of three times for each chunk. Try to make each repetition more accurate than the previous one.

First, let's practice once. Listen and then repeat one time

1. Time one, listen and repeat one time.

2. Time two, listen and repeat one time.

3. Time three, listen and repeat one time.

OK! You are done with the Part A.

Please press the stop button, the yellow square stop button, in Audacity, to stop your recording.

And then save the file as an MP3 file to the desktop. Name your file "your name_A.mp3".

Then you should upload the file to Moodle, for the assignment showing for Friday, May 15, for the self-evaluation task, Part 1.

https://moodle.atlas.uiuc.edu/mod/assignment/index.php?id=249

After saving and uploading Part A, please continue to Part B.

Part B

Now you will repeat the same set of tasks again, using a later portion (segment 2) of your minilecture.

1. First please listen to the following part of your lecture, to familiarize yourself with what you said.

2. This time please listen again to the same recording. I want you to focus on how you used message unit breaks and where you paused and think about how you could pause differently.

3. This time please listen again to the same recording. I want you to focus on how you used primary phrase stress and intonation. Please think about how you could improve these features as you're listening.

4. This time please listen again to the same recording. I want you to focus on how you used word stress and compound noun or compound number stress, if you used those two. Please think about how you could improve these features.

5. This time listen again to the same recording. I want you to focus on your rhythm, including vowel reduction and using clear stress in the right location. Please think about how you could improve these features.

Please click the red record button in Audacity. Do not press the yellow stop button until I tell you to do so.

Next, you will listen to the recording in shorter chunks. After listening to a chunk, repeat, out loud, what you said originally. After speaking, think about your pronunciation accuracy. I'll ask you to listen and repeat a total of three times for each chunk. Try to make each repetition more accurate than the previous one.

1. Time one, listen and repeat one time.

2. Time two, listen and repeat one time.

3. Time three, listen and repeat one time.

[You will repeat the previous three steps for several more chunks.]

OK! You are done with the Part 1 of the self-evaluation for your ML3 recording.

Please press the stop button, the yellow square stop button, in Audacity, to stop your recording.

And then export the file as an MP3 file to your desktop. Name your file "your name_B.mp3"

Then you should upload the file to Moodle, for the assignment showing for Friday, May 15, for the self-evaluation task, Part 1.

https://moodle.atlas.uiuc.edu/mod/assignment/index.php?id=249

After you've saved your file, take a 10-minute break, and then return to the classroom.

After you return to the classroom, you will start the second set of self-evaluation tasks for ML3.

Please go to the next page.

ML3 Self-Evaluation Instructions

2

Part 2

Instructions

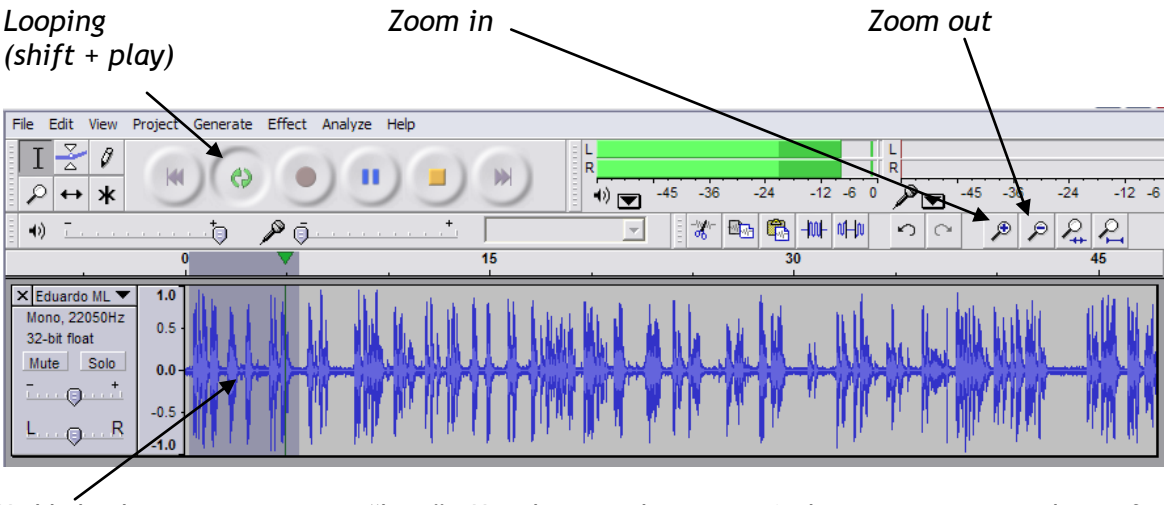
First, open Audacity and size its window to fit the top half of your desktop.

Place your CD in your computer.

Next, use the "File/Open" command in Audacity, to open the .mp3 file on your CD called *"your name_Part_TWO"*.

Practice using the looping and zoom features:

> When you are transcribing, use the "looping" feature, as follows: Use your cursor to highlight 5 to 10 seconds of your recording. Hold down the shift key and then click "play" in Audacity. Doing this will automatically replay the highlighted portion repeatedly, until you click "stop".



Highlight the part you want to "loop". Use the zoom buttons in Audacity to zoom in and out of your recording to make selecting sections easier. Practice using these features now. ***When Sue tells you to do so, continue with the instructions on page 2.***

Instructions

Next you will listen to segments C and D of your ML3 recording.

Part C:

Overview: You will listen to your recording and transcribe it. You will mark the pronunciation features that you used, and then I will ask you to read the transcription and record your voice.

First the ***listening and transcription stage***. Check off each step as you complete it.

___1. In this first step, please listen once to segment C of your lecture, to familiarize yourself with what you said.

___2. Step 2, listen and transcribe all the words and fillers that you said. You may use the looping feature during this step only. Highlight a section; listen as often as needed before moving on to the next section.

During the following steps, you may listen to the recording up to two times. You may pause as often as you need to.

___3. Step 3: After transcribing the words, listen again for message unit (MU) breaks and pauses and mark the actual pauses in your transcription, using a vertical bar or line: | .

___4. Step 4: Listen again for primary phrase stress. If you used it, mark PPS with black dot above the syllable you stressed: •

___5. Step 5: Listen and mark the intonation you used at the end of each message unit. Use these symbols: ↑ (rise), → (rise to midrange), or ↓ (falling).

___6. Step 6: Listen again and mark your use of word and construction stress with an accent mark: ´ .

___7. Listen again and mark the heavy stresses you used on content words, using an open circle above the stressed word: ○

___8. Listen again and mark the following, if you used them:

linking and blending: ‿

trimming of final consonants (e.g., saying "an~~d~~" instead of "and"): cross out trimmed letters

___9. Vowel reduction: Mark unstressed vowels that you did not reduce with a ˘ , e.g., sŏlution.

When you are finished with these transcription tasks, click the "stop" button,  ,
in the Audacity window.

Now the ***recording and oral correction stage***.

Now, use the "File/New" command to open a new window in Audacity. Move this window to the lower half of the screen. You will use this window soon, when it's time to record your voice.

> Overview: You will read and record your transcript 3 times. After each recording, think about your pronunciation accuracy. If you noticed errors in your pronunciation, try to correct these again in the next recording. Try to make each recording more accurate than the previous one.

Now please click the record button in the lower Audacity window and follow these instructions. Do not rehearse before you record your voice in the next steps.

> Say "Time 1". Then read out loud and record your transcript as naturally and accurately as possible in terms of pronunciation.
>
> Say "Time 2". Read out loud and record again, trying to make this production more accurate than the last time.
>
> Say "Time 3". Read out loud and record again, trying to make this production more accurate than the last time.

Please press the stop button,  , in Audacity, to stop your recording. Listen to be sure your voice recorded properly. If it didn't, please check with Sue before continuing.

Save your file to the desktop. Name your file "your name_C.mp3".

To save, click on "File"; "export as MP3". You may be asked to locate the file called "lame_enc.dll". Click "yes" and locate the "lame" file on the desktop. Then save the file as an MP3 file to the desktop.

Next, upload the file to Moodle, for the self-evaluation assignment, Part 2, showing for May 15.

When the file has successfully uploaded to Moodle, close the Audacity window you used to record your voice. When it asks "Save changes before closing?", click "no".

You will repeat these same steps with the next part of your minilecture. The instructions will be the same as above.

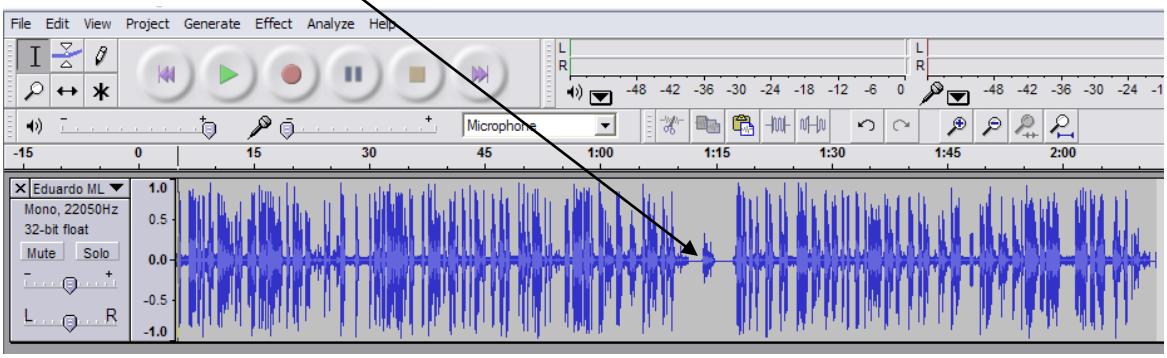
Please turn to the next page when you are ready.

Part D:

Overview: Now you will repeat the same set of tasks again, using a later portion of your minilecture. You will listen to your recording and transcribe it. You will mark the pronunciation features that you used, and then I will ask you to read the transcription and record your voice.

In Audacity, place your cursor at the beginning of the second segment of your ML3 recording. Look for the break between sections of the recording.

Example: Click here:



First the ***listening and transcription stage***.

Check off each step as you complete it.

___1. In this first step, please listen once to segment 2 of your lecture, to familiarize yourself with what you said.

___2. Listen and transcribe all the words and fillers that you said. You may use the looping feature during this step only. Highlight a section; listen as often as needed before moving on to the next section.

During the following steps, you may listen to the recording up to two times. You may pause as often as you need to.

___3. After transcribing the words, listen again for message unit (MU) breaks and pauses and mark the actual pauses in your transcription, using a vertical bar or line: | .

___4. Listen again for primary phrase stress. If you used it, mark PPS with black dot above the syllable you stressed: •

___5. Listen and mark the intonation you used at the end of each message unit. Use these symbols: ↑ (rise), → (rise to midrange), or ↓ (falling).

___6. Listen again and mark word and construction stress with an accent mark: ´ .

___7. Listen again and mark the heavy stresses you used on content words, using an open circle above the stressed word: ○

___8. Listen again and mark the following, if you used them:

linking and blending: ‿

trimming of final consonants (e.g., saying "anɖ" instead of "and"): cross out trimmed letters

___9. Vowel reduction: Mark unstressed vowels that you did not reduce with a ˘ , e.g., sŏlution.

When you are finished with these transcription tasks, click the "stop" button,  , in the top Audacity window.

Now the ***recording and oral correction stage***

Now, use the "File/New" command to open a new window in Audacity. Move this window to the lower half of the screen. You will use this window soon, when it's time to record your voice.

> Overview: You will read and record your transcript 3 times. After each recording, think about your pronunciation accuracy. If you noticed errors in your pronunciation, try to correct these again in the next recording. Try to make each recording more accurate than the previous one.

Now please click the record button in the lower Audacity window and follow these instructions. Do not rehearse before you record your voice in the next steps.

> Say "Time 1". Then read out loud and record your transcript as naturally and accurately as possible in terms of pronunciation.
>
> Say "Time 2". Read out loud and record again, trying to make this production more accurate than the last time.
>
> Say "Time 3". Read out loud and record again, trying to make this production more accurate than the last time.

Please press the stop button,  , in Audacity, to stop your recording. Listen to be sure your voice recorded properly. If it didn't, please check with Sue before continuing.

Save your file to the desktop. Name your file "your name_D.mp3".

To save, click on "File"; "export as MP3". You may be asked to locate the file called "lame_enc.dll". Click "yes" and locate the "lame" file on the desktop. Then save the file as an MP3 file to the desktop.

Next, upload the file to Moodle, for the self-evaluation assignment, Part 2, showing for May 15.

When the file has successfully uploaded to Moodle, close the Audacity window you used to record your voice. When it asks "Save changes before closing?", click "no".

OK! You are done with Part 2 of the self-evaluation for your ML3 recording.

After you've saved your file, take a 10-minute break and then return to the classroom.

Next you will complete the final section of the self-evaluation tasks for ML3.

Please go to the next page.

Name: _____

1. Listen and Transcribe **Checklist**	**SEGMENT C [This page is repeated for segment D]** **Transcribe here:**
___1. Listen only ___2. Listen and transcribe what you really said ___3. Listen and mark all pauses with a | ___4. Listen and mark PPS, with black dot: • ___5. Listen and mark the intonation you used at the end of each message unit: ↑ (rise), → (rise to midrange), or ↓ (falling) ___6. Step 6: Listen again and mark the word stress you used with an accent mark: ´ ___7. Listen and mark heavy stresses you used using ○ above stressed syllables ___8. Mark your use of linking and blending with ‿ Mark trimming of final consonants (e.g., saying "and" instead of "~~and~~"): cross out trimmed letters, if you did so. ___9. Vowel reduction: Mark unstressed vowels that you did <u>not</u> reduce with a ˘ , e.g., sŏlution.	

ML3 Self-Evaluation Instructions

Part 3

3

First, open Audacity and size its window to fit the top half of your desktop.

Place your CD in your computer.

Next, use the "File/Open" command in Audacity, to open the .mp3 file on your CD.

The file to use is labeled with "YOUR NAME_Part_THREE".

You may want to use the looping and zooming features as you complete the listening and transcribing tasks today.

Use the zoom buttons in Audacity to zoom in and out of your recording to make selecting sections easier.

When Sue tells you to do so, continue with the instructions below.

Instructions

In this part, you will listen to segments E and F of your ML3 recording, transcribe them, and mark corrections. Then you will record your voice as you read and correct the segments.

Part E:

First the ***listening and transcription stage***.

◊ Follow the instructions in the "**Listen and Transcribe Checklist"** on the next page.

◊ Check off each step as you complete it.

◊ Use the space on the next page to transcribe your ML3 segment.

◊ When you are finished with the **transcription**, click the "stop" button in the Audacity window.

◊ Then, follow the instructions in the "Mark Corrections Checklist".

◊ After you have completed the two checklists, proceed to the instructions for "recording and oral corrections".

When you are finished with the **transcription** task, click the "stop" button in the Audacity window.

Now the ***recording and oral correction stage***.

Now, use the "File/New" command to open a new window in Audacity. Move this window to the lower half of the screen. You will use this window soon, when it's time to record your voice.

> Overview: You will read and record your transcript 3 times. After each recording, think about your pronunciation accuracy. If you noticed errors in your pronunciation, try to correct these again in the next recording. Try to make each recording more accurate than the previous one.

Now please click the record button in the lower Audacity window and follow these instructions. Do not rehearse before you record your voice in the next steps.

> Say "Time 1". Then read out loud and record your transcript as naturally and accurately as possible in terms of pronunciation.

> Say "Time 2". Read out loud and record again, trying to make this production more accurate than the last time.

> Say "Time 3". Read out loud and record again, trying to make this production more accurate than the last time.

Please press the stop button in Audacity, to stop your recording.

Listen to be sure your voice recorded properly. If it didn't, please check with Sue before continuing.

Name your file using your FirstName_1 (e.g., Sue_1) and save it to the desktop.

Next, upload the file to Moodle, for the self-evaluation assignment showing for today.

When the file has successfully uploaded to Moodle, close the Audacity window you used to record your voice. When it asks "Save changes before closing?", click "no".

Now go to the next page to complete Part F.

Part F:

Now you will work with segment F from your CD file for today.

First the ***listening and transcription stage***.

◊ Follow the instructions in the "**Listen and Transcribe Checklist"** on the next page.

◊ Check off each step as you complete it.

◊ Use the space on the next page to transcribe your ML3 segment.

◊ When you are finished with the **transcription**, click the "stop" button in the Audacity window.

◊ Then, follow the instructions in the "Mark Corrections Checklist".

◊ After you have completed the two checklists, proceed to the instructions for "recording and oral corrections".

Now the ***recording and oral correction stage***.

Now, use the "File/New" command to open a new window in Audacity. Move this window to the lower half of the screen. You will use this window soon, when it's time to record your voice.

> Overview: You will read and record your transcript 3 times. After each recording, think about your pronunciation accuracy. If you noticed errors in your pronunciation, try to correct these again in the next recording. Try to make each recording more accurate than the previous one.

Now please click the record button in the lower Audacity window and follow these instructions. Do not rehearse before you record your voice in the next steps.

> Say "Time 1". Then read out loud and record your transcript as naturally and accurately as possible in terms of pronunciation.

> Say "Time 2". Read out loud and record again, trying to make this production more accurate than the last time.

> Say "Time 3". Read out loud and record again, trying to make this production more accurate than the last time.

Please press the stop button in Audacity, to stop your recording. Listen to be sure your voice recorded properly. If it didn't, please check with Sue before continuing.

Name your file using your FirstName_2 (e.g., Sue_2) and save it to the desktop.

> To save, click on "File"; "export as MP3". You may be asked to locate the file called "lame_enc.dll". Click "yes" and locate the "lame" file on the desktop. Then save the file as an MP3 file to the desktop.

Next, upload the file to Moodle, for the self-evaluation assignment showing for today.

When the file has successfully uploaded to Moodle, close the Audacity window you used to record your voice. When it asks "Save changes before closing?", click "no".

OK! You are done with Part 3 of the self-evaluation for your ML3 recording.

Please return all handouts and your CD to Sue before you leave today. Thanks!

SEGMENT E [This page is repeated for segment F]

1. Listen and Transcribe Checklist	**Transcribe here:**	**2. Mark Corrections Checklist (use a red pencil)**
___1. Listen only ___2. Listen and transcribe what you really said ___3. Listen and mark all pauses with a | ___4. Listen and mark PPS, with black dot: • ___5. Listen and mark the intonation you used at the end of each message unit: ↑ (rise), → (rise to midrange), or ↓ (falling) ___6. Step 6: Listen again and mark the word stress you used with an accent mark: ´ ___7. Listen and mark heavy stresses you used using ○ above stressed syllables ___8. Mark your use of linking and blending with ‿. Mark trimming of final consonants (e.g., saying "~~and~~" instead of "and"): cross out trimmed letters, if you did so. ___9. Vowel reduction: Mark unstressed vowels that you did <u>not</u> reduce with a ˘ , e.g., sŏlution.		___1. Read once to become familiar with the text; make no marks. ___2. Read again and correct MUs; cross out disfluencies ___3. Read again and mark corrections to PPS ___4. Read again and correct intonation ___5. Read again and correct word stress ___6. Read again and correct heavy stresses ___7. Read again and correct linking, blending, and trimming, if they are needed ___9. Review transcription one last time for corrections you missed

Appendix B

Summary of Language Learner Strategy Taxonomies

Source	Categories	Comments
Rubin, 1981 (cited in Grenfell & Macaro, 2007, p. 11)	I. Processes that may contribute directly to learning a. Clarification and verification b. Monitoring c. Memorization d. Guessing/inductive inferencing e. Deductive reasoning f. Practice II. Processes which may contribute indirectly to learning a. Creates opportunities for practice b. Production tasks related to communication	
Naiman et al., 1996, pp. 30-33	<u>Strategies</u> Active task approach Realization of language as a system Realization of language as a means of communication and interaction Management of affective demands Monitoring of L2 performance <u>Techniques</u> (for sound acquisition; included relevant techniques only) Repeating after tapes Reading aloud Listening carefully Talking aloud Practicing different sounds	Condenses Stern (1975)
Oxford, 1989, 1990	Direct memory strategies Direct compensation strategies Direct cognitive strategies Indirect social strategies Indirect metacognitive strategies Indirect affective strategies	

O'Malley & Chamot, 1990	Cognitive Metacognitive Socio-affective	
Cohen, Weaver, Li, 1995	Cognitive: working directly with the learning materials Metacognitive: higher-order processes that have to do with the process of learning Social: seeking L2 interaction for purpose of learning, practice Affective: managing one's emotions in relation to the learning process	These are a combination of Oxford (1990) and Chamot (1987)
Hsiao & Oxford, 2002	Metacognitive strategies Cognitive strategies Memory strategies Social strategies Affective strategies Compensation strategies	They no longer use the direct/indirect distinction
Hassan et al., 2005	Cognitive Metacognitive Socio-affective	This is a summary of O'Malley & Chamot
Macaro, 2006	Cognitive strategies: working memory is directly occupied with cognition about the L2 Metacognitive strategies: conscious evaluation of one's cognitive activities	He provides two main categories of strategies

Appendix C

Mapping of General and Pronunciation LLSs to the Targeted Strategies

	General LLSs				Pronunciation Strategies				
Targeted strategy	Rubin (1981)	Oxford (1990)	O'Malley & Chamot (1990)	Macaro (2006)	Dickerson (1989)	Peterson (2000)	Vitanova & Miller (2002)	Osburne (2003)	Eckstein (2007)
Listening	Direct process: Monitoring	Metacognitive: Self-monitoring Cognitive: Memory strategy: representing sounds in memory Focusing on specific features	Metacognitive: Selective attention Monitoring	Decide to focus attention on targets Remember prior errors for these targets Retain sounds in working memory Compare my production to correct model Apply prior knowledge about pronunciation Does my production of the targeted content sound right? Have I produced each target accurately?	Critical listening Self-monitoring	Self-evaluating Practicing naturalistically Setting goals and objectives Planning for a language task	Self-monitoring Active listening	Focusing on sounds below the syllable-level Focusing on individual syllables Focusing on prosodic structures Monitoring global articulatory gestures Focusing on individual words	Input: Focusing on individual syllables of words Noticing: Noticing the intricate differences between L1 and L2 pronunciation Focusing on supraseg-mentals Intent listening Feedback: Self-monitoring Focusing on supra-segmentals of own speech Hypothesis testing: Rehearsing sounds

General LLSs					Pronunciation Strategies				
Targeted strategy	Rubin (1981)	Oxford (1990)	O'Malley & Chamot (1990)	Macaro (2006)	Dickerson (1989)	Peterson (2000)	Vitanova & Miller (2002)	Osburne (2003)	Eckstein (2007)
Transcription	Direct process: Monitoring	Metacognitive strategy: Self-monitoring Cognitive: Focusing on specific features Analyzing	Cognitive: Note-taking Self-monitoring	Decide to focus attention on targets Remember prior errors for these targets Apply prior knowledge about pronunciation	Not included	Self-evaluating	Self-monitoring Active listening	Focusing on sounds below the syllable-level Focusing on individual syllables Focusing on prosodic structures Focusing on individual words	Input: Focusing on individual syllables of words Noticing: Noticing the intricate differences between L1 and L2 pronunciation Focusing on suprasegmentals

	General LLSs				Pronunciation Strategies				
Targeted strategy	Rubin (1981)	Oxford (1990)	O'Malley & Chamot (1990)	Macaro (2006)	Dickerson (1989)	Peterson (2000)	Vitanova & Miller (2002)	Osburne (2003)	Eckstein (2007)
Annotation	Direct process: Deductive reasoning	Metacognitive: Self-evaluating Cognitive: Highlighting Analyzing and reasoning Predicting Revision	Metacognitive: Selective attention Self-monitoring Self-evaluating Cognitive: Deduction	Decide to focus attention on targets Remember prior errors for these targets Retain sounds in working memory Compare my production to correct model Apply prior knowledge about pronunciation Which production rule applies here? Did my production of the targeted content sound right? Did I produce each target accurately? Decide to revise output	Self-correction Applying rules	Self-evaluating	Self-monitoring	Focusing on individual syllables Focusing on prosodic structures Focusing on individual words	Feedback: Self-monitoring Hypothesis forming: Self-correcting

General LLSs					Pronunciation Strategies				
Targeted strategy	Rubin (1981)	Oxford (1990)	O'Malley & Chamot (1990)	Macaro (2006)	Dickerson (1989)	Peterson (2000)	Vitanova & Miller (2002)	Osburne (2003)	Eckstein (2007)
Rehearsing corrections aloud	Direct process: Practice Indirect: Creates opportunities for practice	Cognitive: Practice Memory strategy representing sounds in memory Focusing on specific features Revision	Metacognitive: Advance preparation Selective attention Self-monitoring Self-evaluating Cognitive: Rehearsal Auditory representation	Decide to focus attention on targets Compare my production to correct model Apply prior knowledge about pronunciation Did my production of the targeted content sound right? Did I produce each target accurately? Decide to revise output	Covert rehearsal	Self-evaluating Planning for a language task Practicing naturalistically	Self-monitoring Active listening	Focusing on local articulatory gesture or single sounds Focusing on sounds below the syllable-level Focusing on individual syllables Focusing on prosodic structures Monitoring global articulatory gestures Focusing on individual words Focusing on memory or imitation	Hypothesis forming: Self-correcting Hypothesis Testing: Rehearsing sounds

Appendix D

Course Timeline and Topics

Day	Date	Topics
1	21-Jan	Placement diagnostic
2 - 3	26-Jan to 28-Jan	Course intro: Syllabus Overview: sounds, rhythm, melody Pronunciation diagnostic assigned
4 - 5	2-Feb to 4-Feb	ML1 One half of students attend each day
6	9-Feb	MUs, rhythm within MUs
7, 8, 9	11-Feb 16-Feb 18-Feb	Self-monitoring tasks: L only LT LTA
10	23-Feb	Articulatory setting Muscle-building activities Vowel review Transcription homework Conferences with instructor Academic vocabulary for pronunciation practice
11	25-Feb	Transcription Compound place names Ash, epsilon, /ey/
12	2-Mar	Alternating rhythm Intro PPS in CW iy/small cap i
13	4-Mar	PPS in FW Theta, eth Tape 2 record GAT/SAT Rhythm in dialogue Stress in street names

Day	Date	Topics
14	9-Mar	Word stress intro -s endings Compound nouns
15	11-Mar	Compound nouns -ed endings Word stress
16	16-Mar	Listening strategy focus: Computer lab listening to NS transcribing NS listen and repeat NS
17	18-Mar	Word stress: first rule
18 to 19	30-Mar 1-Apr	ML2 Individual conferences Tape 3 Sounds MU, rhythm, PPS, word stress
20	6-Apr	PPS contrasts Trimming Linking Word stress ML 2 Self-evaluation
21	8-Apr	Linking l/r Word stress
22	13-Apr	PPS contrasts Yogh, d-yogh Practice annotation (prediction with Matrix dialogue) Homework: listening
23	15-Apr	Sounds: z, d-yogh, yogh Matrix: listening to NS Word stress rule 2
24	20-Apr	Word stress Compound numbers

Day	Date	Topics
25	22-Apr	Intonation, PPS, MUs, listening Assign Tape 4 PPS Sounds Word stress Transcription and annotation Rehearsal
26	27-Apr	Word stress review, add rule 3 Intonation (Q types)
27 - 28	29-Apr 4-May	ML 3 Tape 5 Academic words Word stress Sound level MU, PPS, Inton Question responses (intonation focus) Conferences with instructor
29	6-May	PPS review Question summary Sounds
	11-May 15-May	Self-monitoring tasks in computer lab

Appendix E

Consent Form

Educational Psychology
College of Education
University of Illinois at Urbana-Champaign

Room 226 Education Building, 1310 S. 6th St.
Champaign, IL 61820
phone: 217-333-2245

You are invited to participate in a research study that investigates how speakers of English as a second language develop their pronunciation skills. The study is being carried out by Sue Ingels, TA for ESL ___ and a graduate student in the Department of Educational Psychology at the University of Illinois, under the supervision of Professor Wayne Dickerson, Department of Linguistics.

If you choose to participate in this study, you are agreeing to allow Sue to analyze ESL ___ assignments you will complete for minilectures 1 and 3. The materials to be analyzed are all part of required assignments for ESL ___ and will be completed by you during regular ESL ___ class meetings. The materials include listening, transcribing, and voice recordings for minilectures 1 and 3. During the last week of the semester, you also will fill out a short questionnaire (4 pages) regarding your language-learning experience.

Voluntariness, Risks and Benefits: Participation in this project is completely voluntary. Because Sue is your TA, she will not collect consent forms and will not know if you have agreed to participate until after May 30, 2009, after final grades are submitted. A colleague of Sue's will collect consent forms. You may refuse to participate or discontinue participation at any time. Your decision will not affect your participation in your current or future course of study at UIUC, including ESL ___. If you do change your mind about participating, tell _____ before May 30, 2009, or contact Sue after May 30, 2009. Contact information is given below. For the questionnaire, you may skip any questions you do not wish to answer.

Risks to you are minimal and similar to those encountered in ordinary life. As in any second language setting, you may be sensitive about having someone listen to and evaluate your oral speech. To avoid or minimize the potential for emotional discomfort to you, only the researcher (Sue) and an anonymous rater (also a pronunciation teacher) will have access to your voice recordings and transcriptions. You may request from the researcher a summary of the study's findings, which may be useful for your future English language learning.

Confidentiality: Data from this study will be collected primarily for use in the researcher's doctoral thesis; data also may be included in journal articles and conference presentations. Only the researcher will have access to research results associated with your identity. In the event of publication of this research, no personally identifying information will be disclosed.

Who to Contact with Questions: If you have questions about any part of this study, please ask the researcher, Sue Ingels, whose information is listed below. If you would like a copy of the results of the research, you can contact Sue Ingels. If you have questions about research subjects' rights,

you can contact the research supervisor, Wayne Dickerson, at _____, or the BER Office (217. 333.3023; ber@ed.uiuc.edu).

To submit a consent form or to discontinue participation before May 30, 2009, contact _____; email: ____@illinois.edu; ___ FLB; phone: _____
Researcher: Sue Ingels, Graduate Student, Educational Psychology, University of Illinois email: _____ and phone: _____.

You will receive a copy of this consent form.

I certify that I have read this form and volunteer to participate in this study. I understand that I can withdraw from the study at any time by contacting the researcher or [colleague]. I have been given a copy of this form.

__ I agree to allow my speech to be audio recorded during this study.

__ I agree to allow up to 30 seconds of my voice recording to be used (without personally identifying information) in conference or academic presentations.

Please print name: _____

Signature: _____ Date: _____

Appendix F

Questionnaire

Participant Questionnaire Cover Sheet [insert participant code here]

Your name: _____

Participant Questionnaire [insert participant code here]

Please provide the following information and respond to the following questions. You may skip a question if you do not wish to answer it.

1. Year of birth: _____
2. First language (dialect): _____
3. Years of English language instruction, including elementary, middle and high school, and college (e.g., 1990-96; 2005-2008): _____
4. Date of arrival in the U.S. (month, year): _____
5. Have you lived in another English-speaking country? Circle one: Yes No
6. If yes, where? _____
7. For how long? _____
8. Please list any other languages that you have studied or speak:

Language	Years you studied/learned the language
_____	_____
_____	_____
_____	_____

Please estimate the amount of time spend each day actively interacting in English (circle one):

30 minutes or less 30 to 60 minutes 60 to 90 minutes more than 90 minutes

9. Indicate the types of interaction you usually have in English each week (check all that apply):
___brief comments or responses
___extended conversations or discussions
___meetings with advisor or lab-mates
___conversations with friends or family members
___teaching in English
___participating in classes
___giving presentations
___Other (describe): _____

10. Which UIUC ESL pronunciation classes have you taken? Please list the semester(s) for each (e.g., Spring 2008):

ESL 510 _____ ESL 504 _____ ESL 506 _____

Were you required to take ESL 510? Yes No

11. Have you taken ESL pronunciation courses at another school or university? Yes No

If yes, please indicate the location and year(s) _____

12. Have you worked individually with a pronunciation tutor? If so, please indicate the year and approximate number of times you met: _____

13. Have you taken the SPEAK test? If so, please indicate the dates and scores below:
    First time: date _____ score: _____
    Second time: date _____ score: _____
    Third time: date _____ score: _____

14. Have you completed a SPEAK appeal? If so, please complete the following:
    1ST appeal date (month/year): _____ Outcome: (passed or failed): _____
    2nd appeal date (month/year): _____ Outcome: (passed or failed): _____

15. What was your TOEFL score? _____ Year taken: _____

16. What is your field of study? _____

17. Before taking ESL 504 or other ESL classes at UIUC, have you used any of the following strategies for improving your English pronunciation? If so, please indicate how often, when, and where you used each strategy.

Strategies	How often did you use this strategy?	When and where did you use this strategy?
Listening to a recording of my voice to identify errors:	___ Never ___ 1 to 5 times ___ More than 5 times	Year(s): Location(s):
Transcribing my speech to identify errors:	___ Never ___ 1 to 5 times ___ More than 5 times	Year(s): Location(s):
Orally rehearsing and correcting my errors:	___ Never ___ 1 to 5 times ___ More than 5 times	Year(s): Location(s):

18. Rate each of the following regarding its role in helping you improve your English

(1 = not useful; 5 = very useful):

	not useful				very useful
Listening to a recording of my voice:	1	2	3	4	5
Transcribing my speech:	1	2	3	4	5
Orally rehearsing and correcting my errors:	1	2	3	4	5

19. Do you have any other comments about the listening and transcription tasks you completed today and earlier in the semester?

Thank you for completing this questionnaire!!

Appendix G

Author's Biography

Sue Ann Ingels graduated from Illinois State University in 1981 with a Bachelor of Arts degree in Spanish and psychology. She worked in educational publishing for the next 20 years, including 14 years at Human Kinetics Publishers in Champaign, Illinois. In 2005, she completed a Master of Arts degree in Teaching English as a Second Language from the University of Illinois at Urbana-Champaign and received the Peter Strevens Award for Academic Excellence. While completing her MA and PhD degrees, Sue taught ESL courses in pronunciation, academic and business writing, speaking, and teacher preparation for international teaching assistants. She also coordinated the MATESL teaching practicum and taught pronunciation and special programs for the Intensive English Institute in Urbana, IL. Sue also was a rater for the SPEAK test and an interviewer for that test's replacement, the English Placement Interview, and was a facilitator for Illinois' Graduate Academy for College Teaching. Following completion of her PhD degree, Sue plans to continue teaching, conducting research, and working with second language learners.