

Ingels, S. (2010). The effects of self-monitoring strategy use on the pronunciation of learners of English. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference*, Iowa State University, Sept. 2009. (pp. 67-89), Ames, IA: Iowa State University.

The Effects of Self-Monitoring Strategy Use on the Pronunciation of Learners of English

Sue Ingels

University of Illinois at Urbana-Champaign

This paper describes preliminary results of a classroom-based study on the effectiveness of training advanced second language learners to monitor and correct their non-target use of English suprasegmental features. The independent variables were three levels of self-monitoring [listening only (L), listening + transcription (LT), and listening + transcription + annotation of corrections (LTA)] and rehearsal (R). The dependent variable was an accuracy score for three suprasegmental features: message unit boundaries, primary phrase stress, and intonation. After 16 weeks of pronunciation and self-monitoring instruction, 15 international graduate students at a U.S. university produced a 5-minute mini-lecture, and then used each self-monitoring strategy for a different segment of the lecture. After self-monitoring, participants orally produced the corrected version three times. Rehearsals were recorded and the first and third were analyzed. Preliminary data analysis for seven participants indicates that learner accuracy scores increased following self-monitoring; the greatest improvements occurred after one rehearsal; L was most useful for primary phrase stress accuracy; LTA was best for message unit boundaries and LT for intonation. Accuracy scores improved the most for primary phrase stress, though individual differences were found in learners' performance following use of a particular strategy.

INTRODUCTION

Learning to perceive and produce second language (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 L2 intelligibility and continue to improve their oral skills following the completion of classroom instruction. Improving language skills is an ongoing process and having access to strategies that allow language learners to take control of their own L2 pronunciation learning are considered essential for achieving academic and professional success in L2 contexts (Celce-Murcia, Brinton, & Goodwin, 1996; Dickerson, 1994; Morley, 1991; and others). Thus a primary motivator for this study was the desire to identify effective strategies for improving L2 pronunciation that learners could use independently, without access to specialized resources and without technical linguistic training. Additionally, self-reports from students enrolled in my ESL courses during previous semesters of pronunciation instruction indicated that strategies such as critical listening and transcription helped learners perceive previously unknown suprasegmental features in their own production. As a second language pronunciation teacher, I wanted to gather empirical evidence to determine whether self-monitoring strategies could be used successfully by adult L2 learners.

A variety of factors influence how successful adults are in learning an L2 phonology. Some L2 learners acquire sufficient pronunciation accuracy without explicit instruction (Riney & Flege, 1998). Such learners typically are learning an L2 before puberty or are gifted adult language learners. A number of factors influence 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., 1996; 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 in this study (prospective international teaching assistants, or ITAs, who take ESL pronunciation classes), motivation, daily exposure to English, and prior classroom instruction have not been sufficient for acquiring the necessary accuracy. These learners often are unable to identify when their pronunciation is not target-like and do not notice relevant L2 features in native speaker (NS) speech, though they know their speech differs from target L2 production.

For more than 30 years, language teachers and researchers have been particularly interested in identifying what sets apart “good” language learners from the less successful ones (Oxford, 1990; Rubin, 1975, 1981; Stern, 1975; Vann & Abraham, 1990). A key research goal has been to determine the characteristics and language learning behaviors of successful learners, and the conclusion is that successful language learners often use some of the same strategies to maximize their learning.

Most research has focused on strategies used by learners for improving L2 speaking and general L2 discourse skills (Oxford, 1990). Until recently, only a handful of researchers have focused on specific strategies and techniques that learners can use to improve L2 pronunciation. The most common strategies include noticing and self-correcting (Eckstein, 2007), self-monitoring (Eckstein, 2007; Vitnova & Miller, 2002), and self-evaluating (Peterson, 1997, 2000). 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).

Dickerson (1989) 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 in learning pronunciation and indeed whether such strategies enable learners to make their pronunciation more target-like.

The definition used in this study for language learner strategies (LLS) comes from Hsiao and Oxford (2002, p. 372), who state that LLSs are “active, conscious, purposeful, and attentive,” and from Gu (2007) who describes LLSs as tools learners use to maximize results from their L2 learning efforts. In a pronunciation class that emphasizes student control over learning, students must learn to use strategies for applying pronunciation rules and then apply these strategies in a principled manner during private practice, or “covert rehearsal” (Dickerson, 1989). The process of covert rehearsal includes daily practice, talking aloud in English, monitoring performance for specific features, comparing performance with models stored in memory, making changes in production to match the models, and practicing changes out loud (Dickerson, 2000). In the current study, the use of self-monitoring and self-correcting strategies is based on descriptions provided by Dickerson (2000) and Hahn and Dickerson (1999). Two elements are added here to the range of covert rehearsal techniques: learner use of self-recordings and self-transcription (described in the next section). These latter two were added for the following reasons: a) the strategies had been used for several semesters in our ESL classes, with only anecdotal evidence of their effectiveness; b) they fit the Hsiao and Oxford and Gu definitions of LLSs, as being “active, conscious, purposeful, and attentive” and were skills or “tools” that could be taught to learners as a way to maximize learning; c) very little research existed to indicate whether these strategies were effective for helping ESL learners improve pronunciation; d) if the strategies were indeed useful, they seemed very suitable for use by students in their private study, both during and following completion of classroom instruction (i.e., self-directed study).

Self-Monitoring Strategies Used in This Study and Related Research

Three combinations of self-monitoring strategies are investigated in this study, starting with the most basic strategy (listening only), then adding a second (transcription), and then a third strategy (annotation).

Listening (L). When using this strategy, students listen to short segments of their recorded speech and then attempt to correct any non-target pronunciation. For purposes of this study, the researcher divided student recordings into 1- to 2-message unit segments (typically one or two utterances, or roughly 7 to 14 words), which were presented to students as audio files on a computer. Students did not transcribe their speech nor did they see transcriptions of their speech (though the researcher transcribed all speech samples used in this task in order to determine message unit boundaries prior to setting up the task). Listening (L) required learners to focus on their L2 production at segment, syllable, word, phrase, and discourse levels. A rationale for this strategy is that, when the learner is targeting specific features such as the pitch jump or drop on the syllable receiving primary phrase stress, attention may be drawn to non-target-like pronunciation, thus facilitating bottom-up processing (Izumi, 2003). Research findings indicate that use of critical listening as a strategy for pronunciation improvement results in improved intelligibility (Acton, 1984), reduced use of epenthesis and omission of sounds (Couper, 2003), and production of more accurate word stress and vowel quality (Dickerson, 1987).

Listening + Transcription (LT). When using the LT combination, learners listen critically to their own voice recordings and write down exactly what was spoken, including target-like and non-target-like pronunciation, as well as pauses, restarts, fillers, and repairs. The goal is to create an accurate written record of a speech sample, prior to the evaluation phase. The process of transcribing may allow the learner to notice target forms (Schmidt, 1993) by systematically drawing attention to important L2 pronunciation cues. Research findings indicate that use of transcription may result in improved pronunciation and grammar (Mennim, 2003, 2007) and improved grammar and vocabulary (Lynch, 2007).

Listening + Transcribing + Annotation (LTA). This strategy combination takes the process one step further. Learners refer to a checklist of features to evaluate and systematically review their own transcripts, identify non-target pronunciation, and mark corrections directly on the transcript. Reviewing the transcript is an evaluatory process that learners complete after speaking, when they have the processing resources available to attend to form and its relationship to intended meaning (Trofimovich and Gatbonton, 2006). The assumption in this study is that, because the speaker's message has already been produced, the learner can shift a larger portion of the focus to the correct, corresponding form. No research has been located to date on the role of annotation as a self-monitoring aid.

Rehearsal (R). After performing one of the self-monitoring tasks (L, LT, or LTA) with a given speech excerpt, learners orally produced (rehearsed) each excerpt three times. During and following each rehearsal, learners monitored and evaluated their output, with the goal of identifying modifications that were needed in subsequent rehearsals in order to achieve target-like pronunciation. Research on the effectiveness of task repetition has found evidence for self-correction of pronunciation (Lynch & Maclean, 2001). Repetition and focus on form have been found to exploit familiarity with task, form, and meaning and possibly free up processing resources (Trofimovich & Gatbonton, 2006).

Summary. Except for Couper (2003) and Dickerson (1987), researchers studying self-monitoring strategies have focused on global rather than specific changes in pronunciation features and have looked at use of strategy combinations (e.g., critical listening, transcription, and annotation combined), rather than isolating the effects of each separate strategy component. More research is needed to better understand the value and effectiveness of systematic use of well-defined self-monitoring strategies for improving specific L2 pronunciation features.

Rationale for Selection of Suprasegmental Features for This Study

In the ESL course that is the basis of the study (ESL 504), both segmentals and suprasegmentals are taught, though the latter are emphasized due to students' most typical needs. An emphasis on suprasegmentals also is based on recent research, which 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 Hahn, 1999); absent or

incorrect primary phrase stress (Hahn, 1999, 2004); non-targetlike intonation (Pickering, 2001; Wennerstrom, 1998); and word stress errors (Benrabah, 1997; Guion, Harada, & Clark, 2004; Kawagoe, 2003) all have been shown to negatively impact intelligibility in English. 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 have little or no prior exposure to accented English, so achieving a sufficient level of L2 intelligibility is critical for ITA academic and professional success. Research has not yet 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 listener characteristics, the speaking context and its communicative demands, and the frequency of the L2 speaker's errors.

Seven suprasegmentals were included in the larger study, but the preliminary data reported here relates to three features:

- a) Message unit boundaries (utterances or short phrases separated by brief pauses and/or characteristic intonational patterns)
- b) Primary phrase stress, the most prominent syllable in a message unit, usually occurring on the last content or function word in “new” information in a phrase (“old” or “given” information is not highlighted), but it is also used to signal contrasts, contradictions, comparison, choices, and other information the speaker intends to highlight (Hahn, 1999, 2004). Syllables under PPS usually are longer in duration and higher or lower in pitch, in comparison to surrounding syllables.
- c) Learner use of three categories of intonation were investigated (fall, rise to mid-range, and rise to high range, as described in Hahn & Dickerson, 1999). Falling or “final” intonation is used for signaling completion of a thought; a rise in pitch to mid-range indicates an incomplete (“non-final”) utterance; and a rise to the high end of the pitch range signals certain question types.

MUs are used in this study as the primary unit of spoken discourse. Use of MUs offers a means for choosing equivalent speech samples from each participant. MU length varies within and across speakers, but the other suprasegmental features follow phonological rules that operate within an MU. MUs typically have one prominence (primary phrase stress) and an intonational pattern following the PPS. Because each speaker is producing a different text, speaking at a different rate, and producing varying numbers of morphological units per MU, using the MU as the primary unit of discourse was the best way to standardize the speech samples selected.

RESEARCH QUESTIONS

The purpose of this study was to evaluate the effectiveness of adult L2 learners' use of self-monitoring strategy combinations (*L*, *LT*, *LTA*, and *rehearsing corrections aloud*) to achieve

target-like production of three English suprasegmental features: *message unit boundaries*, *primary phrase stress*, and *intonation* (as defined in Hahn & Dickerson, 1999, and in Cruttenden, 1997). The following research questions were investigated:

1. Effect on Overall Suprasegmental Accuracy

In what ways does strategy use (L, LT, LTA, rehearsal) result in improved pronunciation accuracy? To what extent do strategy combinations have differential effects?

2. Effect on MUs, PPS, Intonation Accuracy

How effective are the strategy combinations for each pronunciation target?

METHOD

Participants

The original group of participants included 15 international graduate students enrolled in a 16-week ESL pronunciation class during Spring 2009. The instructor was the researcher. During the second class meeting, and without the researcher present, a colleague of the researcher collected consent forms from the students who agreed to participate in the study. Participant names remained anonymous to the researcher until after final grades were submitted. All students in the class received the same instruction and completed all of the experimental tasks as part of required coursework. Fifteen of the 16 students in the course participated.

Data from seven native speakers of Mandarin (4 female, 3 male) were analyzed for this paper. All were preparing to re-take the SPEAK test (Speaking Proficiency English Assessment Kit, published by Educational Testing Service), in order to fulfill oral proficiency requirements for teaching assistants. The SPEAK test is a 20-minute oral test, offered once per semester in a computer lab. Students provide timed responses to 12 questions and recordings of their responses are scored by anonymous raters. Rating is holistic, and pronunciation accuracy (intelligibility) is a significant consideration in rating. Students who receive a failing score must take an ESL course or work with a tutor prior to taking the test again. Of the seven participants, 5 had received a SPEAK score of 45, one received 40, and one received a passing score of 50 during the spring 2009 semester, while enrolled in ESL 504.

Participants ranged in age from 23-28 (average = 25 years), all were Mandarin speakers from mainland China, and all had been in the U.S. for 9 – 33 months (average = 16). Average years of instruction was 10 (range = 6 – 13). One student (Vicky) had taken ESL 504 once previously, though with an instructor who did not use the targeted strategies; William and Jeff had taken a

UIUC course (ESL 506) in which pronunciation instruction was a minor component. Jeff had used listening, transcription, and rehearsal five or fewer times during ESL 506; William used listening and rehearsal during three instructor office visits when taking ESL 506. Participants' demographic data are summarized in Table 1.

Table 1. Demographic Data for the Seven Study Participants

Participant	Sex	Age	Home country	Discipline	Months in US	Prior pronunciation instruction?	SPEAK score	Prior strategy use?	Years of English instruction
Andrew	Male	26	China	Computer science	10	no	45 Spring 08	no	13
Jeff	Male	28	China	Statistics	33	UIUC Fall 2008 (minimal focus)	45 Spring 08	All 3 strategies, 5 or fewer times, Fall 2008	11
Nancy	Female	25	China	Statistics	33	no	45 Spring 08	No	6
Vicky	Female	24	China	Sociology	9	Same class Spring 2008	50 Spring 09	No	10
Wendy	Female	23	China	Engineering	9	no	40 Fall 08	Listened to recording 5 or fewer times for TOEFL prep; rehearsal used for prep for presentation	7
William	Male	25	China	Biophysics	9	UIUC Fall 2008 (minimal focus)	45 Spring 08	Listen to recording, rehearsal during office hours with teacher	unknown
Yvonne	Female	24	China	Biology	9	No	45 Fall 08	No	10
Average		25			16				10

Strategy and Pronunciation Instruction

Students received self-monitoring strategy training and instruction on English suprasegmentals throughout the semester (suprasegmental instruction was based on the course text, *Speechcraft* (Hahn & Dickerson, 1999). The instructor provided feedback in several ways: during class (twice weekly sessions of 80 minutes each), as written or audio (recorded) feedback on strategy use and pronunciation following weekly homework assignments, and during individual 20-minute meetings with each student, held three times during the semester (following each of three mini-lecture presentations) to review progress on target features and strategy use.

Materials

Speech samples were gathered from students' third 5-minute mini-lecture (ML), presented during the final week of the semester. Mini-lecture topics represented content from students' fields of study. Students were instructed not to memorize the text nor were they allowed to read from a written text. The last portion of each speech sample typically included the student's responses to audience questions.

Each mini-lecture was audio-recorded in the classroom, using a cordless microphone and a Sony digital recorder. Each mini-lecture was divided by the researcher into six separate but equivalent audio files. These recordings were used by the students to complete the experimental tasks. The MU was chosen as the unit of analysis because the target pronunciation features were suprasegmentals, for which the MU is the relevant environment for analysis. Students in ESL 504 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. Message unit boundaries typically, but do not always, match boundaries of grammatical units such as noun or verb phrases or clauses. Following is an excerpt from a student's mini-lecture, showing the MUs the student produced and also a "target version", or what they would be expected to produce, based on instruction provided during the semester:

ML: so / given the situation that everybody know, / for example if you / if you are / having a / a critical test. /

Target version: so / given the situation that everybody know, / for example / if you are having a critical test. /

Ten days after the final class, students completed the experimental tasks in a lab, using computers equipped with headphones for listening and microphones for recording. Additional materials included written instructions, checklists to guide task completion, and paper and pencils for writing down and annotating (marking up) transcriptions.

Procedure

Due to institutional constraints (i.e., limited class time for completing the tasks and the need for participant identities to remain anonymous to the researcher), experimental tasks were completed on one day, during a 2 to 2-1/2 hour session with regular breaks. Tasks were completed in the following order: L, LT, LTA. For each participant, early, middle, and late speech segments were randomly assigned to the three self-monitoring tasks. This was done to prevent systematic bias based on location of a segment in the 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 the L task, participants listened five times to a segment representing approximately one sixth of their lecture (15 MUs). They were instructed to listen to one suprasegmental feature each time. Next, they were presented aurally with a 1- to 2-MU portion of the larger segment and were told to “Listen and repeat one time”. This listening and repetition (rehearsal) phase occurred three times for each portion until each 15-MU segment was completed. Each rehearsal was audio-recorded. This process was repeated for the next 15-MU segment from a later portion of the mini-lecture.

For the LT task, participants followed a similar process of listening to two different segments of their lecture. During the first listening, participants transcribed the segment. During subsequent listenings, they focused separately on each targeted suprasegmental feature.¹ After completing the transcription, participants read and orally corrected the transcribed segment a total of three times. Each reading was audio-recorded.

For the LTA task, participants completed the listening and transcription steps as described for the LT task (using two new sections of the lecture), then systematically reviewed the transcriptions for non-target pronunciation, marked corrections (annotated) in a different colored pencil, and then read and orally corrected each segment a total of three times. Each reading was audio-recorded. Figure 1 summarizes the experimental procedures.

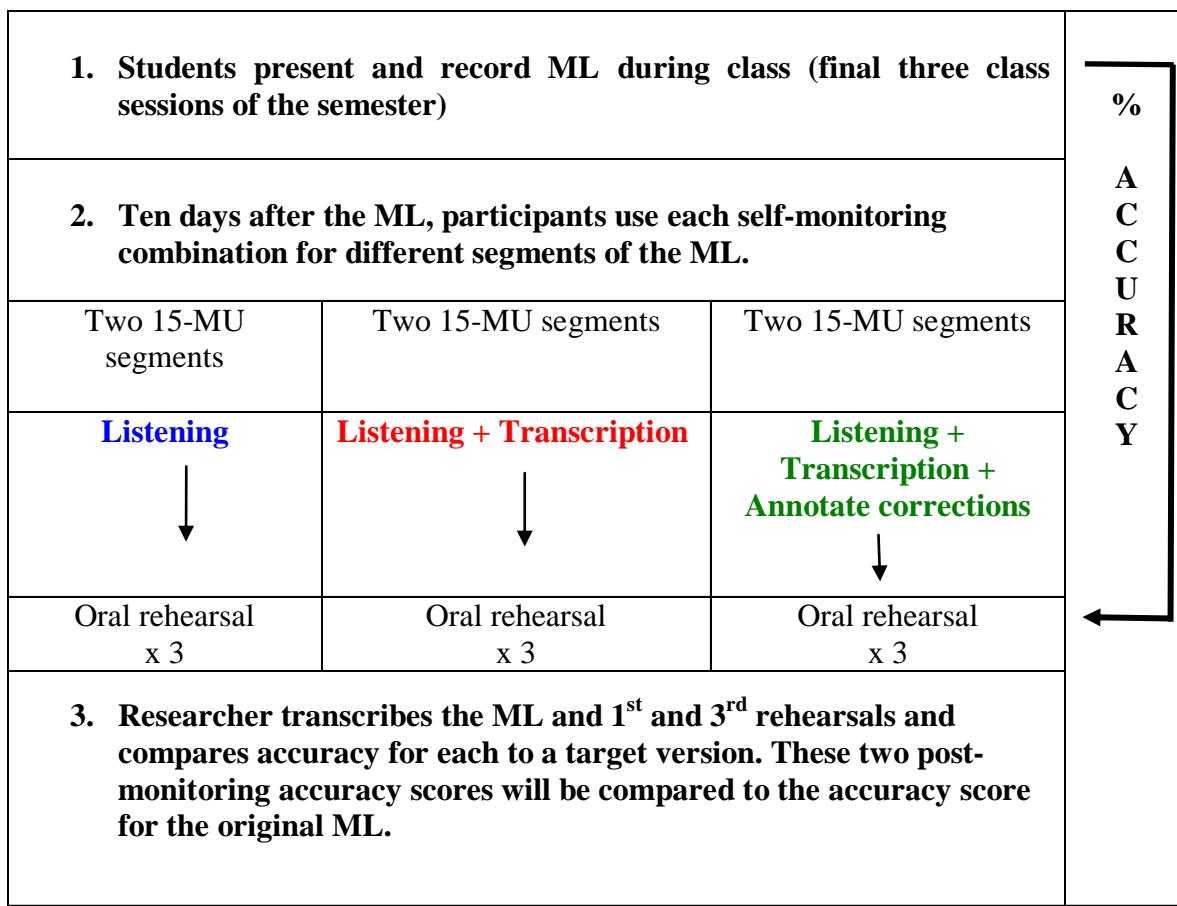


Figure 1. Summary of experimental procedures.

Preparation of data for analysis

Each participant's mini-lecture recording was transcribed by the researcher, and MU boundaries, PPS, and intonation patterns were noted. The researcher prepared a "target" version of the mini-lecture text (Figure 2), showing the suprasegmental features the students would be expected to produce correctly, based on the semester's instruction. The researcher transcribed participants' first and third oral rehearsals, marking the targeted suprasegmental features. The following transcription conventions were used:

MU boundaries are marked with a forward slash (/).

PPS is marked using a solid black dot (●) above the syllable receiving PPS.

When a syllable that should be unstressed is given a heavy stress (but not PPS), an open circle (○) is placed above the syllable. This often occurred on pronouns that should have been unstressed. Intonation patterns were noted as follows: a comma (,) denotes a rise to mid- or

high-range (non-final or question intonation); a period (.) denotes a fall to low range, or phrase-final intonation.

Target version:

we will make decision, / whether I like it, / or dislike it. /

Mini-lecture:

we will / make decision. / whether I like it. / or dislike it. /

1st Rehearsal:

we will make decision, / whether / we like it, / or dislike it. /

↑ ↑ ↑ ↑
1 2 3 4

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

Data Analysis

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., mini-lecture 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, that 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 (from an original 30 MUs).

Accuracy scores. During data analysis, MU boundaries, PPS, and intonation were marked as correct or incorrect, based on a comparison with what the student could be expected to do following the course’s instruction. The correct targets were tallied to result in an accuracy score (represented as a percentage of correct targets) for each strategy combination used. The accuracy score for the mini-lecture was considered the baseline level for each participant (what they could do without self-monitoring). Accuracy scores for the first oral rehearsal were used to determine the extent a specific strategy combination resulted in increased accuracy for each pronunciation target. Scores for the third oral rehearsal provided evidence for the effectiveness of repeated rehearsals.

Accuracy scores were calculated for each of the following data categories: (a) by strategy combination for all three suprasegmentals combined: for the total group, and separately for each individual; (b) by strategy combination for each suprasegmental feature: for the group and for each individual.

RESULTS

Research question 1: In what ways does strategy use (L, LT, LTA, rehearsal) result in improved pronunciation accuracy? To what extent do strategy combinations have differential effects?

Group Results. Overall group results (Table 2) show that accuracy scores for the three pronunciation targets improved for each type of self-monitoring, from 76% to 82% for L, 76% to 81% for LT, and from 75% to 83% for LTA. The sample size did not permit tests of statistical significance for these differences. Grouping the data masks individual differences in strategy effectiveness, thus individual results are provided next.

Table 2. Group Accuracy Scores and Percentage Difference Following the Use of Self-Monitoring and Rehearsal

All Participants	Mini-lecture 3	1 st Rehearsal	% difference
Across all conditions	76%	82%	+6%
Listening only	76%	82%	+6%
Listening + Transcription	76%	81%	+5%
Listening + Transcription + Annotation	75%	83%	+8%

Individual results. Individual results for each strategy combination, regardless of pronunciation target, appear in Figure 3. L was the most effective strategy for three participants (Nancy, 11%; Vicky, 7%; Wendy, 13%) and was also highly effective for a fourth (Yvonne, 11%); LT was most effective for only one (Andrew, 11%); and LTA was highest for William (11%) and Yvonne (14%), and LT and LTA were both most effective for Jeff (10%). Thus L and LTA resulted in the most frequent improvement. These individual findings are consistent with the group results showing greater effects for LTA and L as compared to LT.

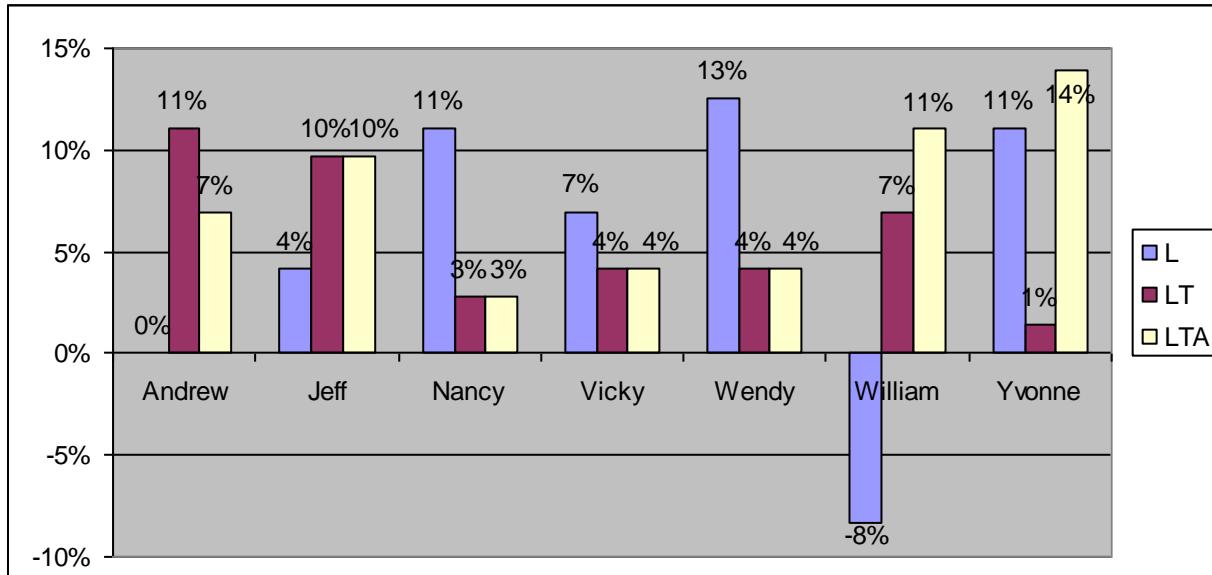


Figure 3. Percentage improvement for each participant, by strategy combination.

Effects of Rehearsal. When looking at group average results, regardless of strategy type, the greatest improvement following self-monitoring occurred with the first rehearsal, with a difference of 7% for MUs and PPS, and 6% for Intonation (Table 3, Figure 4). As expected, accuracy gains were not as strong between the first and third rehearsal; there simply is less room for improvement following subsequent rehearsals. However, when looking at each strategy type, we see that accuracy following the third rehearsal was the same as or lower than R1 in several instances, including use of LTA for message units, L and LT for PPS, and LT and LTA for Intonation.

Table 3. Effects of Rehearsal

	Message units			Primary Phrase Stress			Intonation		
	ML	R1	R3	ML	R1	R3	ML	R1	R3
L only	82%	86%	88%	65%	75%	74%	85%	87%	88%
LT	80%	86%	90%	65%	68%	68%	81%	89%	86%
LTA	83%	92%	90%	61%	67%	70%	81%	88%	85%
Average	81%	88%	89%	63%	70%	71%	82%	88%	86%

Note: R1 = first rehearsal; R3 = third rehearsal

Because the accuracy score gains were greatest for the first rehearsal and scores changed minimally or slightly declined at R3, the remaining data analysis focused on the change from the mini-lecture to the 1st rehearsal.

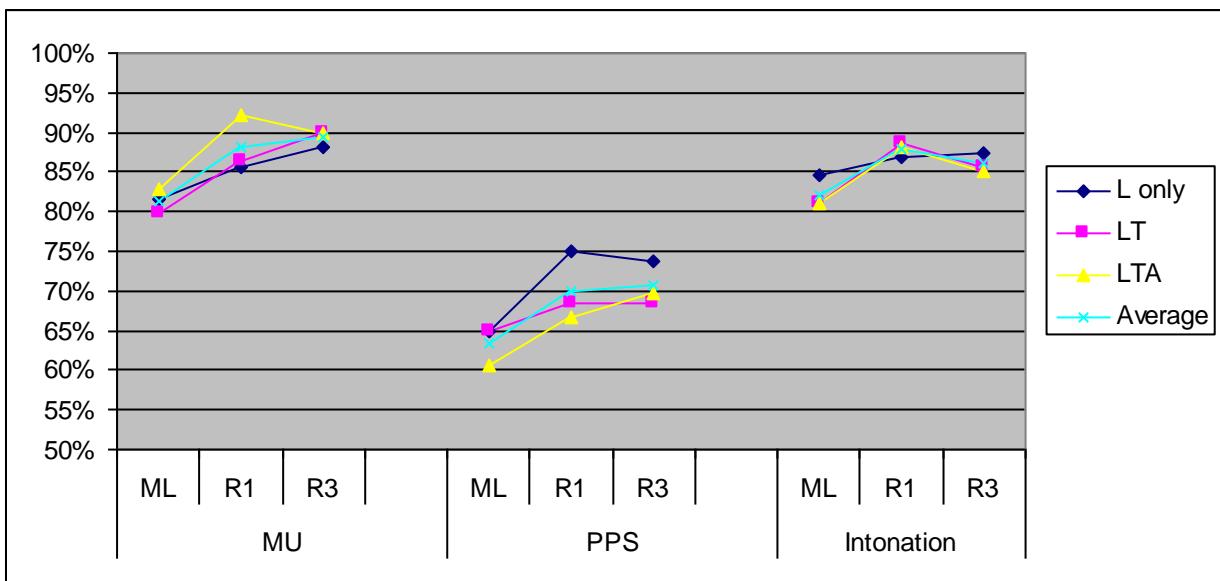


Figure 4. Change in accuracy scores (group values) for the mini-lecture (ML), 1st rehearsal (R1), and 3rd rehearsal (R3), for each pronunciation target, by strategy combination.

Research question 2. Effect on MUs, PPS, Intonation Accuracy - How effective are the strategy combinations for each pronunciation target?

When looking at group results by pronunciation target (Table 4), the difference in accuracy from the mini-lecture to the first rehearsal was greatest for PPS (7%), followed by MUs (6.7%), and then Intonation (6%). These findings pattern with the beginning accuracy score for each target, with PPS the lowest (65%), followed by MUs (82%), and Intonation starting at the highest accuracy level (85%). Thus PPS had the greatest room for improvement.

Table 4. Group Values for Percentage Difference Between Accuracy at the Mini-Lecture and the 1st Rehearsal, by Target.

% difference	
PPS	7.0%
MUs	6.7%
Intonation	6.0%

When looking at each pronunciation target, results indicate accuracy increases were greatest for MUs when using LTA (9.4%), followed by LT (6.4%), and L (4.3%). The pattern differed for PPS, with L (10.4%) > LTA (6%) > LT (4.4%). And for Intonation, yet a different pattern emerged: LT (8%) > LTA (7.3%) > L (2.6%) (See Figure 5).

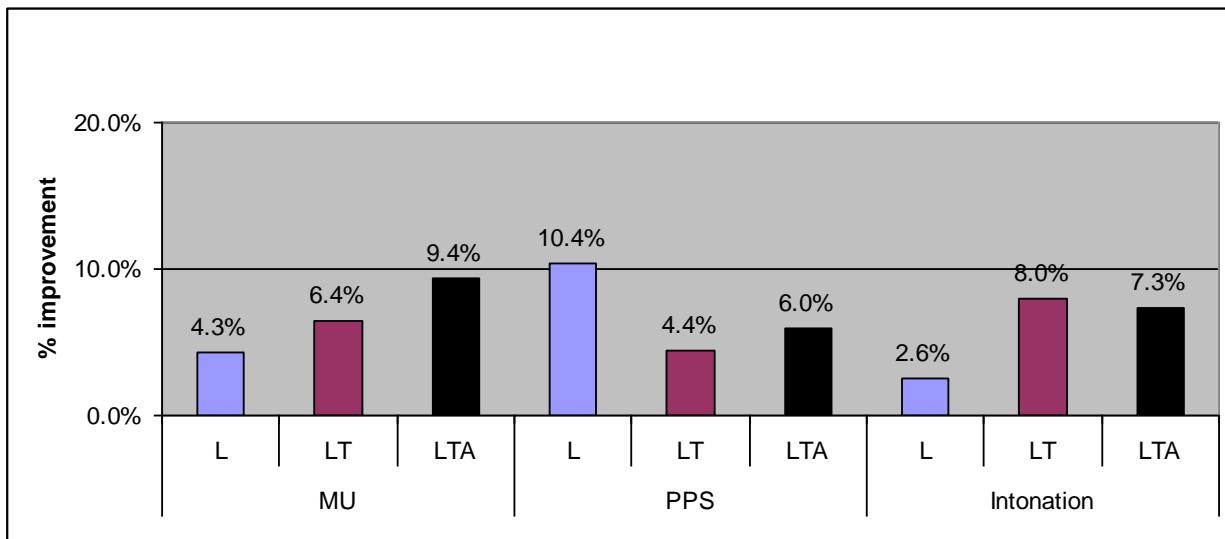


Figure 5. Percentage improvement from mini-lecture to 1st rehearsal for the group, by pronunciation target and strategy combination.

Individual Results for Pronunciation Targets. Individual performance mirrors group results for each target. Use of one of the transcription combinations (LTA or LT) was beneficial for most participants, with the exception of PPS. LTA was most successful for MUs. L was the most useful strategy for PPS, and LT was most useful for Intonation (Table 5).

Table 5. Breakdown of Effective Strategy Combinations by Target, by Participant. (Totals May Exceed 7.)

Target	Strategy Combination		
	L	LT	LTA
MUs	3	2	5
PPS	5	2	1
Intonation	1	4	3

DISCUSSION

Overall, the use of self-monitoring strategies results in increased pronunciation accuracy scores and specific strategy combinations have different effects for each of the three pronunciation targets.

Research question 1: In what ways does strategy use (L, LT, LTA, rehearsal) result in improved pronunciation accuracy? To what extent do strategy combinations have differential effects?

Group results suggest that the LTA combination is most effective for this group of L2 learners (Table 2). These findings are consistent with those of Mennim (2003, 2007) and suggest that the

use of transcription may help learners notice, focus on, and identify pronunciation features that are more difficult to detect when listening holistically to one's speech (Izumi, 2003; Schmidt, 1993). Pennington and Ellis (2001) also concluded that unless L2 speakers' attention is drawn to the functioning of prosody in the L2, they likely won't attend to it, especially when the role of prosody is different in their L1 (as is the case with Mandarin). Thus training is necessary to help L2 learners, even at an advanced stage, to notice and interpret suprasegmental features.

Following are other factors important in the interpretation of the data.

Prior Use of Strategies. Three participants indicated prior use of self-monitoring strategies. Jeff used all three combinations during another UIUC ESL course, most likely following three presentations. Wendy used listening and rehearsal for preparation for a presentation and for the TOEFL test. William indicated that he used listening and rehearsal during office hour visits with his instructor (typically 3 to 4 times per semester). In each case the strategies were used 5 or fewer times, which makes it unlikely that these participants received sufficient practice to influence their performance in this study.

Possibility of Ceiling Effects. Beginning accuracy levels for several participants exceeded 90% for MUs (Jeff, William, and Wendy) and Intonation (Nancy, William), suggesting that these participants had nearly mastered these features and had little room for gains in accuracy. Thus these features were not the best choice for this study for these participants. However, the story for PPS is different. All participants' initial accuracy scores for PPS started at or below 83%, with the lowest level at 42% (Figure 4), so ceiling effects are not likely operating for this pronunciation target. Analysis of other target features in the remaining speech data, including linking, vowel reduction, and word and compound noun stress, may, like PPS, start with lower accuracy levels and thus offer clearer insights into strategy effects. L2 instructors will want to rely on results of diagnostic tests to determine the targets for which learners will derive the greatest benefit from self-monitoring.

Effects of Task Load. An unexpected finding from this study was that multiple rehearsals did not necessarily result in improved accuracy scores. Group results showed that differences in R1 and R3 accuracy were only +1% for L and LT and -2% for LTA. Several factors may help explain this pattern.

- **Fatigue.** Participants completed the tasks during one session of 2.5 hours or less. This was done to accommodate the course schedule, which limited the amount of class time available for completing the experimental tasks. 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, when looking at individual performance across the three tasks, accuracy levels for only 2 of the 5 participants declined over time (from L to LTA) and only for Intonation (Nancy and Wendy). Though I did not track the

amount of time that each strategy combination required, I observed that participants took longest to complete the Listening only task. Except for PPS, this strategy was least effective for most learners, thus some participants may have felt boredom or frustration over the length of the task.

- **Lack of motivation to complete the task accurately.** Most students in ESL 504 failed the SPEAK test and are in the class because it is a requirement to retake the test. Thus one might expect students' intrinsic motivation levels to be low. Motivation levels were not directly targeted in this study. However, on a post-experiment questionnaire, six of seven participants rated the usefulness of Listening and Rehearsing higher than transcription. This might suggest that participants should perform better on subsequent rehearsals regardless of task, though the results do not support this.
- **Task effects.** Across the three rehearsals in the LT and LTA tasks, learners may have started to rush and focus less on producing accurate target features during R2 and R3. The findings from the current study conflict with those of Lynch and McLean (2001), who found that repetition resulted in pronunciation improvements, due to task familiarity. Further investigation is needed to better understand the current findings.
- **Memory effects.** Learners may be over-taxing short-term and working memory as they process multiple chunks of text during rehearsal.
- **Second guessing.** By consciously attending to L2 features that they previously produced automatically, participants may be doubting their initial choices and "correcting" something that was accurate to begin with (Willingham & Goedert-Eschmann, 1999). If second-guessing is a factor, one would expect greater declines on MUs and Intonation, given their high beginning accuracy levels. As noted previously, this did happen for Intonation for Nancy and Wendy, but not for other participants.

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 maximize benefits from rehearsal and how to maintain concentration on the task.

Individual Differences. Participant performance was variable across the three strategy combinations (Figure 3). For example, Andrew was most successful using LT and showed no change in accuracy when using L. Conversely, Nancy, Vicky, and Wendy were most successful using L. These seemingly inconsistent results at the individual level may be due to differences in learning styles (e.g., visual vs. aural styles), second-guessing previous decisions, or incorrect

application of rules. Participant interviews and analysis of transcripts from the LT and LTA tasks may provide insights into possible learner preferences for a particular strategy combination.

Research question 2. Effect on MUs, PPS, Intonation Accuracy - How effective are the strategy combinations for each pronunciation target?

Results suggest that learners may find greater success when using LTA for MUs, LT for Intonation, and L for PPS (Figure 5). What are possible reasons for these differences? First, participants had already achieved a high level of accuracy for MUs and Intonation and thus may be better equipped to monitor these targets successfully. Perhaps MU boundaries and intonation contours are easier to “see” in a transcription, and harder to remember when listening only. That is, the visual cues of slashes (/) for MU boundaries and arrows for intonation markings (↑↓) may be more memorable and easier to implement during rehearsal. Perhaps when reading, learners do not need to rely on memory and can use the visual cues of the transcript as reminders. Regarding PPS, perhaps PPS is more salient for learners when presented *aurally* rather than *visually*. Seeing the dot (●) over a syllable receiving PPS may not be sufficiently meaningful, but, during listening and speaking, learners may be able to hear and feel the difference between target-like and non-target-like PPS. As noted earlier, at least for PPS, analysis of a transcript may lead to more second-guessing than listening only. Further analysis of learner transcriptions and comparisons of their predictions and performance in the LTA condition are needed to better answer this question.

Limitations

Several limitations of this study are evident and offer potential for future research. First, the sample size is small. These preliminary findings will be expanded as the data for the remaining eight participants are analyzed. Second, additional data are needed to determine the efficacy of self-monitoring for learners at low and intermediate levels. Third, inconsistencies in task types may influence outcomes. Some tasks required reading, others listening only. The original mini-lecture was presented to an audience, however, the experimental tasks were not. Though the focus of this study is on strategies for self-study, an interesting question is to what extent interaction may result in more accurate production when using self-monitoring.² Further work is needed to understand how such task differences may affect the study’s results. Fourth, R2 was not analyzed, making it unclear how accuracy scores evolved from R1 to R3. R2 effects will be investigated during the next data analysis stage. Fifth, the results don’t offer insights into long-term effects of self-monitoring. And lastly, the tasks used in this study were not authentic 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.

CONCLUSIONS

The results of this study move us a few steps further in our understanding of the merits of strategy use for L2 pronunciation improvement. Prior research has looked at global pronunciation change (e.g., Acton, 1984; Mennim, 2003) and has not tried to isolate the effects of self-monitoring strategy use on message unit boundaries, primary phrase stress, and intonation accuracy. 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.

Completion of the data analysis for this study will further illuminate the efficacy of self-monitoring strategies for L2 pronunciation and the relationship between strategy types and pronunciation targets. Further research is needed to explore long-term benefits of self-monitoring strategy use and further expand our understanding of how language teachers can facilitate L2 pronunciation improvement for their students.

NOTES

1. The full study will analyze the participants' use of rhythm facilitators and word and construction stress. These targets are not included in this paper.
2. I would like to thank Colleen Meyers for pointing out the potential value of interaction in making this study's pedagogical tasks more realistic and the possibility that the communicative component may further enhance self-monitoring accuracy.

ABOUT THE AUTHOR

Sue Ingels, MATESL, is a PhD candidate in the department of Educational Psychology at the University of Illinois at Urbana-Champaign. Her ESL teaching experience includes courses in pronunciation, business communication, academic and business writing, and international teaching assistant training, all at UIUC. She also has coordinated the MATESL teaching practicum. Her research interests include acquisition of English suprasegmentals by ESL learners, L2 phonological acquisition and pronunciation instruction, and teacher education.
Email: ingels@illinois.edu

REFERENCES

- Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly*, 18(1), 71-85.
- Benrabah, M. (1997). Word-stress: A source of unintelligibility in English. *International Review of Applied Linguistics*, 35, 157-165.
-

- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
- Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect*, 18(3), 53-70
- Cruttenden, A. (1997). *Intonation* (2nd ed.). Cambridge, UK: Cambridge University Press.
- 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 Publications.
- 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. (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, Canada, March 18, 2000.
- 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.
- Gu, P. Y. (2007). Foreword. In A. D. Cohen & E. Macaro (Eds.), *Language learner strategies* (pp. vii-viii). Oxford: Oxford University Press.
- 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. (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, MI: University of Michigan.

- 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.
- 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.
- Juffs, A., & Rodriquez, G. (2007). Working memory in context: Differential effects on comprehension of relative clauses and anaphor binding. Presented at Second Language Research Forum, Urbana, IL, October 13, 2007.
- 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.
- Lightbown, P. M., & Spada, N. (2006). *How languages are learned* (3rd ed.). Oxford: Oxford University.
- 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.
- 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* 63(2), 81-97.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481-520.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal*, 84, 372-389.

- 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.
- Peterson, S. S. (2000). *Pronunciation learning strategies: A first look*. Unpublished research report. (ERIC Document Reproduction Service ED 450 599; FL 026 618).
- Pickering, L. (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly*, 35(2), 233-255.
- 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.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.
- Stern, H. H. (1975). What can we learn from the good language learner? *The Canadian Modern Language Review*, 31(4), 304-318.
- 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.
- Tyler, A., Jeffries, A., & Davies, C. (1988). The effect of discourse structuring devices on listener perceptions of coherence in non-native university teachers' spoken discourse. *World Englishes*, 7(2), 101-110.
- Vann, R. J., & Abraham, R. G. (1990). Strategies of unsuccessful language learners. *TESOL Quarterly*, 24(2), 177-198.
- Vitnova, G., & Miller, A. (2002). Reflective practice in pronunciation learning. *The Internet TESL Journal*, VIII(1), April 24, 2008.
- Wennerstrom, A. (1998). Intonation as cohesion in academic discourse. *Studies in Second Language Acquisition*, 20(1), 1-25.
- Willingham, D. B., & Goedert-Eschmann, K. (1999). The relation between implicit and explicit learning: Evidence for parallel development. *Psychological Science*, (10)6, 531-534.