Pearson, P., Pickering, L., & Da Silva, R. (2011). The impact of computer assisted pronunciation training on the improvement of Vietnamese learner production of English syllable margins. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 169-180), Ames, IA: Iowa State University.

THE IMPACT OF COMPUTER ASSISTED PRONUNCIATION TRAINING ON THE IMPROVEMENT OF VIETNAMESE LEARNER PRODUCTION OF ENGLISH SYLLABLE MARGINS

Pamela Pearson, Georgia State University

Lucy Pickering, Texas A & M, College Station

Rachel Da Silva, Georgia State University

In this study we investigate the utilization of computer assisted pronunciation training to the issue of syllable margin production by Vietnamese learners of English. Thirteen intermediate Vietnamese students were recruited from an intensive English program. A pre-test was administered to elicit student performance on the production of syllable margins (e.g. onsets and codas). Treatment materials comprised a second set of items featuring audio files recorded by a native speaker of American English and used as the target models for training. Eight 30-minute tutoring sessions were conducted in which students used a KayPentax Computerized Speech Laboratory to compare their self-produced spectrograms with the prerecorded target spectrograms. Post-test measures indicated some significant movement toward more target-like production of syllable margins. This paper describes the pedagogical methods and materials used in this study with an emphasis on word list creation, design of intervention and coding and analysis. We outline the benefits of this training for pronunciation improvement including the positive reactions of both teachers and students to the use of this teaching technique.

Among practitioners of ESL/EFL, certain pronunciation challenges are recognized as in some way intractable. Not that they cannot be improved, but that improvement may be rare and comes at a high cost in terms of time and effort for both learners and teachers. One such area is the production of syllable margins in the speech of Vietnamese learners (Hansen, 2006). One relatively recent approach to addressing obstinate pronunciation issues is the use of multimodal or visual feedback through the use of speech visualization technology (Hardison, 2004, 2005; Levis & Pickering, 2004). Computer assisted pronunciation training (CAPT) (Levis, 2007) encompasses a range of tools and techniques which include the use of spectrograms, pitch contours and statistical information on acoustic features of oral production. What they have in common is an *analytic-linguistic* approach to pronunciation training as opposed to the more traditional *imitative-intuitive* approach (Celce-Murcia et al, 1996). Initial investigation of these techniques suggests that they may enhance auditory-only instruction thus we conducted a pilot study to examine the effect of CAPT on the syllable margin production in Vietnamese language learners.

LITERATURE REVIEW

Syllable structure transfer in L2 acquisition has been frequently documented in the SLA literature (see Leather & James, 1991 for a review). Although the investigation of the phonological development of Vietnamese learners of English (VLEs) is rare, early studies (Benson, 1988; Sato, 1984, 1985) emphasize the role of L1 transfer in syllable margin production leading to consonant cluster reduction and the differential production of word-initial and word-final clusters. This is reiterated by Osburne (1996) in her case study of a Vietnamese learner. In two interviews over a span of six years, she demonstrates that cluster reduction is not random but rather regular and "influenced subtly by the expectations of L1 syllable structure" (Osburne, 1996, p. 175). Vietnamese lacks consonant clusters (Avery & Ehrlich, 1992; Honey, 1987) and has a tendency toward open syllables, which may contribute to word ending deletion (Hwa-Froelich et al., 2002; Osburne, 1996).

In addition to the investigation of linguistic constraints including L1 transfer, Hansen (2006) examines socio-affective factors including social identity, gender, context of use, and investment on the acquisition of syllable onsets and codas of a Vietnamese married couple who begin as newcomers to the US. During her 10-month study, Hansen finds little overall change in terms of ultimate acquisition (approximately 85% for onsets and 40% for codas), however she documents some emerging production modifications toward targetlike production that may lead to change over a longer period of time. These studies present a somewhat gloomy picture for VLEs as traditional methods of instruction do not appear to have a significant impact on development. Syllable structure may become a "stabilized" aspect of Vietnamese interlanguage phonology (Han, 2004; Long, 2003).

CAPT techniques in pronunciation instruction represent a novel approach only recently widely accessible to teachers and students through downloadable freeware such as *Praat* (Boersma & Weenink, 2005). Speech visualization feedback has been utilized to examine segmental and suprasegmental issues in pronunciation (Hardison, 2004; Hirata, 2004; Lambacher, 1999; Levis & Pickering, 2004). Lambacher (1999) uses "electronic visual feedback" with Japanese learners to practice the production of English consonant contrasts: oral and nasal stops, liquids, and fricatives. Although he claims that the students made "articulatory adjustments" based on direct comparison with the L1 models (p. 145), he provides no empirical evidence as to the actual acquisition of these contrasts. Hardison (2005) used visual displays of learners' pitch contours to provide contextualized prosody training and found significant improvement following training in addition to positive reactions from learners regarding the use of visual feedback. To our knowledge, no one has used this kind of feedback to explore the acquisition of English syllable margins.

In light of the intelligibility problems displayed by our VLEs, many of whom reported having studied English for five years or more, we undertook the following study to investigate if we could destabilize VLE syllable structure through an approach designed to cultivate apperceived input (Gass & Selinker, 2001). In this form of CAPT, spectrograms of the learners' spoken production were used to provide visual feedback.

METHOD

Participants

In total, 13 students participated in the study: five (4 = F; 1 = M) during the summer of 2008, and eight (2 = F; 6 = M) in the spring of 2009. The mean age of the participants was 23 (range: 19 to 28 years); the mean amount of time they had spent studying English in the US was five months (range: 0.5 months to one year); and the mean length of time they had spent studying English in Vietnam was five years (range: two years to eight years). At the time of the intervention, each student was enrolled in an Intensive English Program (IEP) at a major university in the Southeastern United States. The students were taking 15 semester hours of classes at the IEP including a three-hour course in Oral Communication. Participant proficiency level was classified as pre-intermediate (n = 5), intermediate (n = 6), or upper intermediate (n = 2) based on the results of an in-house placement exam, which included a proficiency interview.

Materials

Pre- and post-tests were constructed to elicit the participants' performance with regard to the perception and production of syllable margins (i.e., onsets and codas.) Each test comprised (a) 30 multiple-choice items to test perception, (b) 55 items to test production, and (c) a semi-structured speaking task[1]. To avoid any potential interference caused by suprasegmental or literacy issues, all items were monosyllabic words. Decisions regarding which syllable margins to incorporate were made according to their appearance in the lower tiers of accuracy percentiles on similar tasks with VLEs, as reported by Hansen (2006, pp. 66, 78-79).

Our final word list comprised 10 singleton onsets, 10 two-member onsets, and five three-member onsets (the maximum possible in English); it also contained 10 items in each of the following categories: single, double and triple codas (see Appendix A for complete listing). To ensure internal consistency and equivalence of forms, items on the pretest were repeated but randomized on the post-test.

A second set of items featuring the same onsets and codas was created for use in the actual treatment in order to minimize the practice effect (Appendix A). From this second set, the materials required to administer the treatment were developed. WAV audio files for each item were digitally recorded by one of the researchers, a native speaker of American English, and filed for use as spectrogram models, and a packet of student-use materials depicting items in both print and image form was created (Appendix B).

Procedures

The treatment was carried out in either instructor-pair or instructor-small group (n = 3) tutoring sessions using a Kay Pentax Computerized Speech Laboratory (CSL) Model 4500 equipped with Real-Time Spectrogram software. Eight 30-minute sessions (two per week for four weeks) were conducted. The first and last meetings were reserved for testing, and there were six tutoring sessions in between. A breakdown of the session activities schedule follows in Table 1.

[1] These data are not examined here

Table 1. *Session Activity Summary*

Week	Session	Activity	Session	Activity
1	1	Intro to the study & CSL; IRB consent; pretest	2	Single onsets
2	3	Double onsets	4	Triple onsets
3	5	Single codas	6	Double codas
4	7	Triple codas	8	Post-test

During each session, students were provided with speech visualization feedback using the CSL (see Figure 1 below). Students were shown a model of the target word and trained to recognize the components of the spectrogram. They then produced the word and analyzed their self-produced spectrograms in light of the model spectrograms. Sessions were driven by an emphasis on mastery, meaning that students were allowed to record items as many times as necessary in order to produce spectrograms with which they were satisfied.

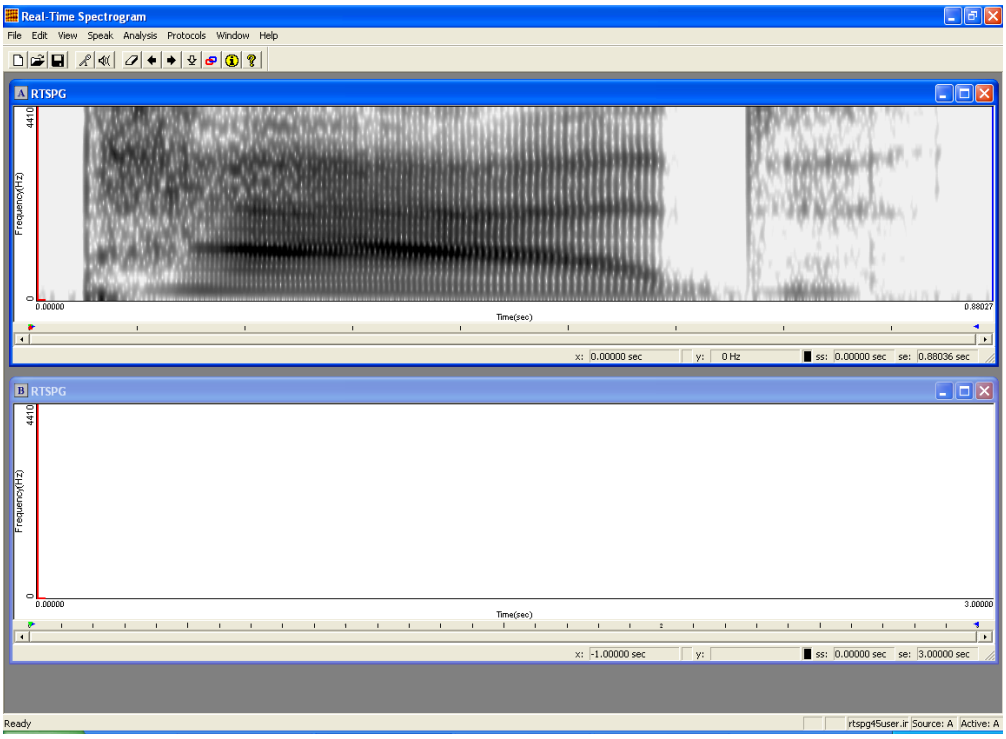

Figure 1. Sample screenshot of *pad* in Real-Time Spectrogram software

Data Analysis

In order to determine whether positive gains occurred from Time 1 to Time 2, paired samples *t*-tests were run in *SPSS 15.0* (SPSS Corporation, 2006). Perception data was entered as either 'correct' or 'incorrect'. However, due to a high group mean on the pretest, and consequent lack of a normal distribution of the scores, a *t*-test was not a viable analysis to perform for the measure of perception and simple percentages were used instead. Production data were coded prior to analysis. Coding forms were created for both versions of the test (see Appendix C) which allowed for both the dichotomous coding of items – targetlike or non-targetlike – and also for the diagnosis of non-targetlike productions according to categories used by Hansen (2006): deletion, substitution, epenthesis and other. In order to receive a coding of 'targetlike', an item had to be judged as such by both raters. Using Cohen's Kappa, interrater reliability across the entire data set was .71. In simple percentages, inter-rater reliability was 88.7% (agreement = 634/715 items) for the pretest, and 82.5% (agreement = 590/715) for the post-test data.

RESULTS

With regard to the perception tests, the participants collectively scored 338 out of a total of 390 multiple-choice items resulting in an 86.7% rate of accuracy on the pretest. As a group, they performed marginally higher on the post-test, making 36 errors on the 390 items resulting in a 90.8% rate of accuracy with an overall rate of improvement of 4.1%. In contrast to the high rate of accuracy on the perception pretest, on the production pretest, the participants produced only 255 targetlike syllable margins (out of a possible 715) for a combined accuracy rate of 35.7%. On the post-test measure, however, students demonstrated a considerably higher rate of accuracy. Overall, the rate of accuracy rose to 52.2% (373 targetlike productions on 715 items). A paired samples *t*-test conducted with this data found a statistically significant increase of group means (from 19.6 on the pretest to 28.7) ($p = .01$ where $p < .05$).

Items were also clustered by syllable margin categories and ranked by percentage of targetlike production: single onsets (80%), double onsets (62%), triple onsets (44%), single codas (42%), double codas (40%), and triple codas (28%). The data show that codas proved to be more troublesome for the VLEs than onsets in both perception and production tests. These results are highly consistent with the findings of Hansen (2006).

DISCUSSION

This study suggests that pedagogical intervention may be valuable in terms of improving VLEs' production of syllable margins at least in the short term. Following tutoring sessions employing CAPT methods, early post-tests show a significant difference in overall accuracy of production of forms produced in a word-list format. In post-test production, raters reported more difficulty in assessing targetlike vs. non-targetlike forms resulting in a slightly lower agreement rate. We suggest that these changes in production show that syllable structure can be "destabilized" in the IL of these learners. Clearly, it will be important to expand the treatment period, vary tasks, and allow for continued delayed post-testing in our follow up work in order to determine the applicability and duration of the effect.

In addition, in further tests we will reconsider some of our word list choices. Four of the items were produced in a non-targetlike manner by all five of the participants at both Times 1 and 2: *fudge*, *teethe*, *cloths*, and *minced*. We realized that word recognition (e.g., *cloths* might look like the more frequent *clothes*) could play an important role in whether the students were able to achieve targetlike production (Venkatagiri & Levis, 2007). In the future, we will control for frequency by consulting a word frequency index such as the Compleat Lexical Tutor (CLT) (Cobb, 2008), which compares items to the 2000 most common word families in the Brown Corpus of American English. Of the 55 items on our word list, only 15 appear on the CLT list: *bagged*, *bands*, *bath*, *cloths*, *hid*, *hits*, *jump*, *locked*, *log*, *pray*, *pumped*, *swim*, *tap*, and *thanks*.

Finally, with regard to pedagogical implications, CAPT provided both teachers and students with an alternative form of feedback to address this particular intractable error. Through the visual aspect of the spectrogram-enhanced feedback, students were able to notice a gap in their production which was something they had not been able to do through the imitative-intuitive approach alone. Toward the end of the study, one participant commented, "I didn't know I don't say the endings of words. So, now I know, and I'm trying it when I speak in English." The teachers also responded positively to the technique. They perceived the results of the CAPT approach to be successful, and they believed they had found a more effective way to provide feedback to this particular group of learners. One teacher noted that "what [she] liked most about the CSL is that it takes the personal element out of the feedback. Instead of telling students 'no' and making them repeat over and over again, we were instead able to give them a positive goal to work towards." Although we can only report these additional findings anecdotally at this stage, we plan to include a more formal investigation of motivational benefits in our further work in this area.

APPENDIX A: Word Lists

Test items* (onsets)	Treatment items (onsets)	Test items (codas)	Treatment items (codas)
1. pat /p/	pad	26. tap /p/	lap
2. ban /b/	bat	27. hid /d/	kid
3. ten /t/	tech	28. lash /ʃ/	bash
4. zoo /z/	zoom	29. leave /v/	weave
5. shot /ʃ/	shop	30. pull /l/	full
6. vain /v/	vase	31. fudge /dʒ/	budge
7. jug /dʒ/	just	32. bath /θ/	math
8. chip /tʃ/	chick	33. teethe /ð/	breathe
9. there /ð/	they	34. locked /kt/	socked
10. think /θ/	thing	35. bagged /gd/	tagged
11. blue /bl/	bloom	36. wished /ʃt/	fished
12. play /pl/	plane	37. wind /nd/	sinned
13. crew /kɹ/	cruise	38. puffs /fs/	huffs
14. pray /pɹ/	pry	39. hits /ts/	mitts
15. swim /sw/	swish	40. cloths /θs/	moths
16. class /kl/	clam	41. pierce /ɹs/	fierce
17. three /θɹ/	throw	42. jump /mp/	bump
18. small /sm/	smog	43. clasp /sp/	asp
19. frog /fɹ/	free	44. pumped /mpt/	jumped
20. tweed /tw/	tweak	45. minced /nst/	winced
21. splash /spl/	split	46. lunched /tʃt/	bunched
22. street /stɹ/	string	47. faxed /kst/	taxed
23. spray /spɹ/	spread	48. helped /lpt/	yelped
24. scream /skɹ/	screen	49. turns /ɹnz/	burns
25. squid /skw/	square	50. flasks /sks/	masks
26. log /g/	dog	51. months /nθs/	ninths
27. robe /b/	lobe	52. bands /ndz/	hands
		53. thanks /Nks/	banks

*Accompanied by IPA transcriptions of phonological targets

APPENDIX B: Sample of Teaching Materials

APPENDIX C: Pretest production coding form sample

Name: _____

Rater: _____

Pretest Coding: Speaking

Onsets

1.**p**at	TL	non-TL: deletion, substitution, epenthesis, other
2.**b**an	TL	non-TL: deletion, substitution, epenthesis, other
3.**t**en	TL	non-TL: deletion, substitution, epenthesis, other
4.**z**oo	TL	non-TL: deletion, substitution, epenthesis, other
5.**sh**ot	TL	non-TL: deletion, substitution, epenthesis, other
6.**v**ain	TL	non-TL: deletion, substitution, epenthesis, other
7.**j**ug	TL	non-TL: deletion, substitution, epenthesis, other
8.**ch**ip	TL	non-TL: deletion, substitution, epenthesis, other
9.**th**ere	TL	non-TL: deletion, substitution, epenthesis, other
10. **th**ink	TL	non-TL: deletion, substitution, epenthesis, other

REFERENCES

Avery , P. W. & Ehrlich, S. (1992). *Teaching American pronunciation*. Oxford, UK: Oxford University Press.

Benson, B. (1988). Universal preference for open syllables as an independent process in interlanguage. phonology. *Language Learning* 38:221-235.

Boersma, P., & Weenink, D. (2005). Praat: doing phonetics by computer (Version 4.3.01) [Computer program]. Retrieved from http://www.praat.org/

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

Cobb, T. (2008). *The Compleat Lexical Tutor* [online computer software]. Retrieved from http://132.208.224.131/

Gass, S. M. & Selinker, L. (2001). *Second language acquisition: An introductory course (2nd edition)*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Han, Z. (2004). Fossilization: Five central issues. *International Journal of Applied Linguistics*,14(2), 212-242.

Hansen, J. G. (2006). *Acquiring a non-native phonology: Linguistic constraints and social barriers.* London, UK: Continuum.

Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning and Technology, 8*(1), pp. 34-52.

Hardison, D. M. (2005). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal* 22: 175-190.

Hirata, Y. (2004). Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning*, *17*(3-4), 357-376.

Honey, P. J. (1987). Vietnamese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems (1st edition)*. Cambridge, UK: Cambridge University Press.

Hwa-Froelich, D., Hodson, B. W., & Edwards, H. T. (2002). Characteristics of Vietnamese phonology. *American Journal of Speech – Language Pathology, 11*(3), 264-273.

Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals.* Oxford, UK: Oxford University Press.

Lambacher, S. (1999). A CALL tool for improving second language acquisition of English consonants by Japanese learners. *Computer Assisted Language Learning*, *12*(2), 137-56.

Leather, J. & James, A. (1991). The acquisition of second language speech. *Studies in Second Language Acquisition* 13: 305-41.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27,* 1-19.

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

Long, M. H. (2003). Stabilization and fossilization in interlanguage development. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 487-535). Oxford, UK: Blackwell Publishing.

Osburne, A. G. (1996). Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics, 17*(2), 164-181.

Sato, C. (1984). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning* 34: 43-57.

Sato, C. (1985). Task variation in interlanguage phonology. In S. Gass & C. Madden (eds). *Input in second language acquisition.* Cambridge, MA: Newbury House.

SPSS Corporation (2006). *SPSS 15.0* [computer software]. Chicago, IL: SPSS Inc.

Venkatagiri, H. S. & Levis, J. M. (2007). Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness*, 16(4), 263-277.

STUDENTS' AWARENESS OF SPANISH SPIRANTIZATION ALLOPHONIC RULE

Manuela González-Bueno, University of Kansas

Marcela Quintana-Lara, Universidad Arturo Prat

El propósito de este estudio es determinar si aprendices de español como lengua extranjera con diferentes niveles de habilidad lingüística tienen conciencia de la regla de espirantización de las oclusivas sonoras del español. Se espera que los resultados ayuden a determinar si esta regla se puede adquirir de forma natural, sin intervención pedagógica alguna, o si, por el contrario, se necesita entrenamiento enfocado en la forma para hacer que los estudiantes sean conscientes de la regla y consecuentemente la aprendan.

The purpose of this study is to determine the extent to which L2 Spanish learners at different levels of proficiency are aware of the spirantization rule of Spanish voiced stops. The results of this study help determine whether general exposure to the language is enough to acquire the target rule or more form-focused phonological training is needed to bring their attention to the rule and consequently learn it.

INTRODUCTION