

Web-based pronunciation training: an outline of materials and their effectiveness

著者(英)	Midori IBA
journal or publication title	Language and Culture : The Journal of the Institute for Language and Culture
volume	10
page range	29-40
year	2006-03-15
URL	http://doi.org/10.14990/00000428

Web-based Pronunciation Training: An Outline of Materials and their Effectiveness

Midori IBA

Abstract

Studies on the sound system currently emphasize the importance of suprasegmental features such as rhythm, intonation and stress. Many researchers argue that because of their major role in communication, these features merit greater priority than attention to individual sounds in the curriculum. This doesn't mean, of course, that work on individual vowels and consonants is irrelevant, but these features should not be the starting point or the pedagogical focus when learning another language. Yet, the traditional approach to teaching English pronunciation gives priority to individual sounds, and teachers and texts have devoted the majority of time and effort to English vowels first, then the detailed coverage of the vowel chart and diphthongs, followed by work on consonants. As this segmental work takes up most of the time available, only brief attention to the prosodic features of English is paid. Since it was hard to find a textbook starting with prosodic features, I created an original website for practicing English pronunciation with the financial support of my university in 2003. This paper explains the main features of the site, reports on its actual use in class, and examines the effect the software has on developing the prosodic features and improving the accuracy of individual English sounds of my students. Reference is made to the computerized sound analysis that gauged whether there had been any improvement in the prosodic skills of the students after practicing English pronunciation with the site.

Introduction

For ESL/EFL learners to achieve successful communication, prosodic features are as critical as individual sounds because they turn the basic building blocks of the sound systems into words, utterances, and discourse.

With suprasegmentals and connected speech, however, the misunderstanding is apt to be of a more serious nature. Learners who use incorrect rhythm patterns or who do not connect words together are at best frustrating to the native-speaking listener; more seriously, if these learners use improper intonation contours, they can be perceived as abrupt or even rude; and if the stress and rhythm patterns are too non-native like, the speakers who produce them may not be understood at all (Celce-Muria,

Brinton, & Goodwin 1996)

In recent years, the field of pronunciation teaching appears to have adopted a more balanced viewpoint with regard to the importance of both the segmental and suprasegmental aspects of language (Derwing, Munro, & Wiebe, 1998). However, stress, rhythm, and adjustments in connected speech can be easily overlooked in the actual language classroom. Every year at the beginning of my class, I ask students whether they had enough time in other English classes to study these features and I always receive mostly negative answers. In fact, it seems that even segmental features are not sufficiently covered.

In the 2003 academic school year, I decided to create a website for practicing English pronunciation and applied to my university for financial support. The plan was accepted and I was introduced to a software company for technical assistance in developing the site. The software was uploaded and I started to use it in my intermediate-level listening course in 2004.¹⁾

In this article, the main features of the software are outlined in the first section, and its application to the classroom is dealt with in the following section. Next, the results of a study on the effectiveness of the computer-assisted pronunciation training are reported, using data collected from the students and analyzed by computer.

1. The main features of the software

1-1. Priority of prosodic features

An ideal training tool should be one that can produce a significant improvement in both the segmental and suprasegmental levels of the spoken language. With this in mind, I designed the software as follows.

There are three parts to the software. Part 1 consists of five units explaining prosodic features and sound changes (i.e., assimilation, linking and elision). Computerized visual displays of pitch contours and sound waves are provided in this part. In Part 2, typical English consonant sounds which are considered problematic for Japanese learners are explained with videotaped mouth movements of NS models. Part 3 deals with vowel sounds of English and Japanese. The units within this part focus on explaining the difference between similar vowels with the visual mouth movements of NS models as in Part 2. Each unit in the three parts has a practice section.

Part 1	Part 2	Part 3
1-1 Rhythm	2-1 /r/ and /l/	3-1 /i/-/ɪ/ and 「い」
1-2 Intonation	2-2 /f/ and /h/	3-2 /e/-/ɛ/ and 「え」
1-3 Contraction	2-3 /b/ and /v/	3-3 /æ/-/ə/ and 「あ」
1-4 Assimilation	2-4 /θ/ and /s/	3-4 /o/-/ɔ/ and 「お」
1-5 Linking	2-5 /ð/ and /z/	3-5 /u/-/ʊ/ and 「う」
1-6 Elision	2-6 /θ/ and /s/ 2-7 /n/ and /ŋ/	

Table 1. Website contents

The reason for first dealing with rhythm is that it reflects a hierarchical organization of the temporal sequence of speech sounds into syllables and higher level units of prosodic and syntactic structure. Part of learning a spoken language is the acquisition of its systematic rhythmic organization. Every language has a characteristic rhythm, and listeners expect to hear speakers use that rhythm. It is vital that learners use the rhythm that is characteristic of the target language, in this case English. There must be an alternation of stressed and unstressed syllables, with stressed syllables occurring mostly on a regular beat, which is quite different from the rhythm of the Japanese language. If Japanese learners of English are not aware of differences such as these, they will have difficulty in being understood. Intonation is also important for intelligibility because it is used to express intentions. An inappropriate use of intonation patterns may lead to misunderstanding, just as mispronounced words can. Furthermore, misunderstandings may result in judgements concerning the speaker's attitude and behavior. If he constantly uses low pitch without much melodic variation, listeners might form the impression that he is, for example, a low-key person. Regarding sound changes in connected speech in English, most Japanese learners are not aware of phenomena such as contraction, assimilation, linking and elision. The purpose of Part 1 in the software is, therefore, to make Japanese learners of English sufficiently aware of the difference in the prosody of the two languages and help them build on this basic awareness.

1-2. Two models of English pronunciation

I chose a GA speaker and an RP speaker as the models for English pronunciation so that a learner can choose one (or both) of those traditional native speaker accents. I am not a strong supporter of 'standard English,' nor have I ever assumed that students have approximating native speaker speech patterns as their aim. Especially in a world where there are four-times as many non-native speakers of English as native speakers, it is surely wrong for a teacher to push students to feel that anything other than these two accents is an imperfection. More importantly, accents are a fundamental part of identi-

ty. Students might feel uncomfortable if they are forced to imitate a particular accent of English. Even teachers, regardless of whether they are native or non-native speakers, might feel the same, especially in the case that their own accents are not GA or RP. Indeed, teaching pronunciation is a delicate issue, but if students are learning English as a means of communication, it would seem obvious that the ultimate goal must surely be to approximate the speech sounds of native speakers of the target language. In that case, the ‘target’ sounds should be shown to students. Some students might aspire to a native speaker accent, while others will not be interested in their accent at all. Teachers should assure learners that both approaches are perfectly legitimate in learning pronunciation. The reason why I chose the two major accents for models is simply because they are accepted as ‘major,’ ‘standard,’ or sometimes ‘prestigious’ in the real world. Of course there are many controversial aspects to this the issue, but I prefer to argue for a neutral position when it comes to choosing a model for English pronunciation.

2. Application of the software to a language classroom

2-1. About the learners

The software has been uploaded and two classes of students who take intermediate listening courses at Konan University use the website. The number of the students is 52. Most of them are in their second year.

2-2. How is the website used in class?

Since there is no class of English pronunciation at the university, pronunciation is taught somewhat differently by individual teachers. In my case, in the first semester this year, students completed 10 pronunciation training sessions (See Table 1, Part 1-1 Rhythm to Part 2-4 /θ/ and /s/). As the training is one part of a listening class, each session lasts only about 10 minutes and it is not connected with the material in the textbooks they use in class. From the student’s perspective, pronunciation training is most likely seen as an additional task at the end of the listening class.

2-3. Assignments

After practicing a new unit, students were asked to review the unit at least twice a week outside the classroom, using their own computer or computers that are available at the university.

3. Examining the effectiveness of the training

3-1. Design of the study

The purpose of the current study is to gauge whether there had been any improvement in the prosodic skills of the students after practicing English pronunciation with the site. Segmental accuracy such as consonant duration was also investigated along with the prosody.

A pretest/posttest design was used to measure the effects of the 10 weeks of training (10 sessions of about 10 minutes each) on English prosody using computerized visual displays of pitch contours. In the training sessions, the pronunciation of eight English consonants (/r/, /l/, /f/, /h/, /b/, /v/, /θ/, and /s/) was also covered by using a visual display of mouth movements. For the pretest, participants were asked to produce a set of new sentences to test the generalizability of the training. They were unaware that their production would also be rated for segmental accuracy as well as for prosody.

3-2. Method

Participants. A total of 43 native-speaking Japanese students participated in this study. 38 participants were female and five were male; all were undergraduates at Konan University in the first semester of their second or third year. None had studied or lived abroad at that point. Through preliminary interviews, I concluded that the participants were representative of the second year college level of proficiency (low intermediate). All participants were offered the opportunity to obtain feedback on testing performance when data analysis was completed.

Materials

1. Headset: SONY HS-90
2. Computer: COMPAQ EVO D320ST/CTC1,7
3. Amplifier: SONY Educational Amplifier EA-80
4. Speech Analyzer: Sugi Speech Analyzer

The selection of sentences for testing and training followed these guidelines: a) familiar vocabulary, b) functional value to college students, c) sustained phonation to provide the best possible continuous display of pitch contour, and d) a range of sounds including those that are often difficult for Japanese speakers. For feedback purposes, the training sentences were recorded by two NSs. Each NS was instructed to look at a sentence printed on a list and then produce it at a conversational rate of speech into the microphone. Sentences were played back to check the intelligibility and naturalness of expression. Testing sentences were also recorded by NSs and stored on hard disk for

later comparison purposes.

Procedure. Participants were tested and trained individually. For the pretest at the beginning of the listening course and the posttest after the ten training sessions, they were shown an identical passage²⁾ printed on a card. No pronunciation instruction was given. They were allowed to practice the sentences aloud before recording. After looking at each sentence, they were instructed to produce the sentences at a conversational rate into the microphone. The sentences were recorded on their own MD and then stored as separate files on hard disk.

For the training sessions, participants practiced English pronunciation with the website as explained in 2-2.

3-3. Results: Comparison of the two recordings

Because the size of the collected data was so large, I chose one part of the passage (“How strange I thought”) spoken by each student for the sound analysis. The data were analyzed by a speech analyzer to measure the duration of the consonant /s/ in “strange,” and to compare pitch contours of the pretest recording and posttest recording with NS pitch contours. The reason for measuring the duration is that the English consonant /s/ needs more power in pronunciation and its duration is longer than when pronounced in Japanese.

The following are the results of a comparison of the two recordings, before-practice (BP) and after-practice (AP) (N=43).

The duration of /s/ Most students’ (42 out of 43) pronunciation of /s/ was longer after training. See Table 2.

Tuesday class

Student	(BP) Duration of /s/ ms	(BP) Duration of ex. s	(AP) Duration of /s/ ms	(AP) Duration of ex. ms
T1	104	1356	109	1232
T2	130	1550	146	1454
T3	113	1656	92	1470
T4	78	1470	117	1448
T5	83	1992	116	1556
T6	140	1667	147	1416
T7	124	2036	135	1388
T8	102	1366	106	1129
T9	99	1561	145	1383
T10	202	1473	207	1295
T11	103	1788	120	1764
T12	135	1348	142	1325
T13	88	1577	115	1343
T14	138	1400	145	1255
T15	113	1521	136	1269
T16	83	1070	125	1292
T17	57	1002	105	1150
T18	149	1547	165	1821
T19	99	1218	173	1411

Friday class

Student	(BP) Duration of /s/ ms	(BP) Duration of ex. ms	(AP) Duration of /s/ ms	(AP) Duration of ex. ms
F1	107	1160	127	1433
F2	71	1235	138	1283
F3	147	1903	157	1684
F4	84	1178	93	1251
F5	136	1471	144	1407
F6	95	1235	132	1188
F7	126	1313	135	1121
F8	106	1312	133	1389
F9	97	1375	137	1341
F10	119	1378	123	1211
F11	131	1340	148	1243
F12	116	1143	154	1226
F13	134	1427	151	1197
F14	82	1412	127	1398
F15	136	1441	149	1370
F16	139	1545	154	1536
F17	117	1888	130	1627
F18	128	1582	144	1366
F19	104	1133	143	1206
F20	91	1207	129	1209
F21	98	1567	116	1375
F22	83	1441	119	1277
F23	80	1197	122	1183
F24	105	1798	120	1153

Table 2. Duration of the consonant /s/ for each student Before Practice (BP) and After Practice (AP). Duration of the expression (ex.) is also shown.

The duration of the expression “How strange, I thought” 32 students spoke faster after training (See Table 2).

Pitch. In the comparison of pitch contours, much improvement was found in all cases. Regarding auditory impressions, most of the cases can be recognized as ‘improved.’ The following are examples from some participants.

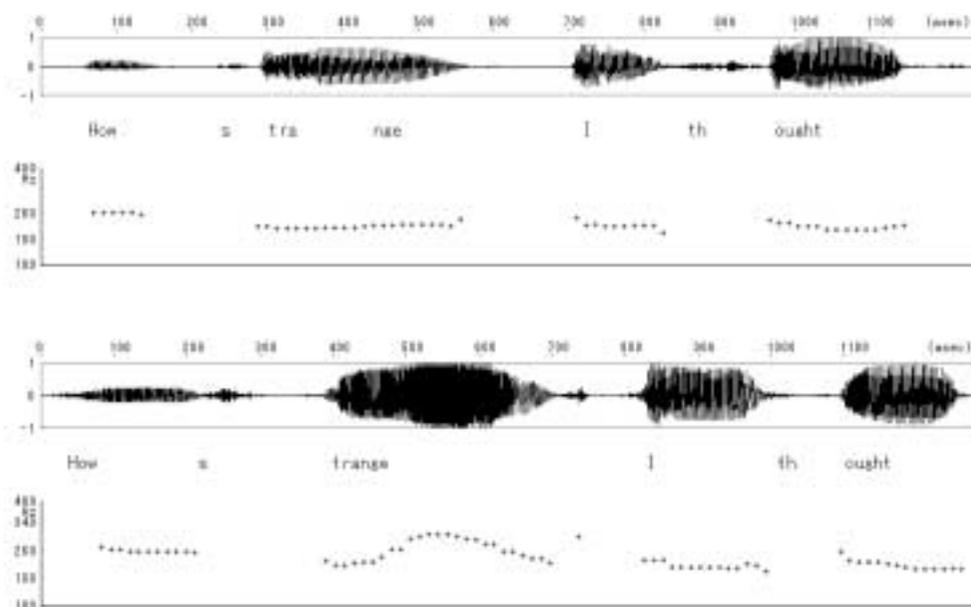


Figure 1. Above: pitch contours of a student (F2 in Table 2) before practice. Below: pitch contours of the same student after practice.

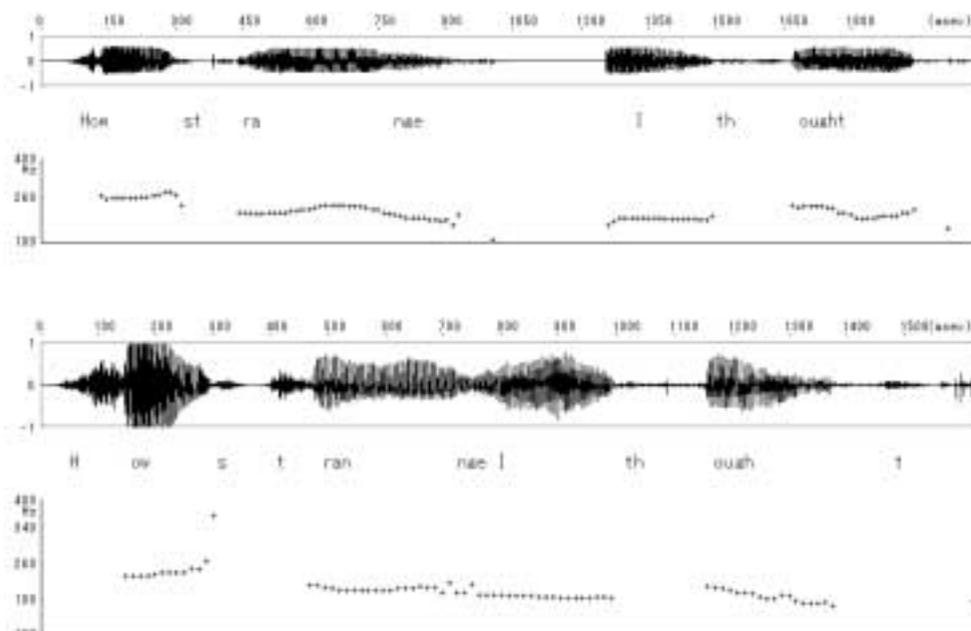


Figure 2. Above: pitch contours of a student (T5 in Table 2) before practice. Below: pitch contours of the same student after practice.

In Figure 1, there is not much variation in pitch and pitch contours look flat before practice. The change in pitch after practice shows that the speaker's utterance sounds more melodic than before. In Figure 2, the pitch contours still look flat after practice, but the auditory impression is quite different from BP. The AP utterance sounds more natural as English because its rhythm has changed to near-stress-timed. The stress-timed rhythm of the English utterance with related obscuration of weak syllables is the prime distinguishing feature of the language's pronunciation. A comparison of all the pitch contours of BP and AP of the participants as in Figure 1 and Figure 2 shows significant effects of the training in the acquisition of L2 prosody. However, students' pitch contours AP are still different from those of the NS models (see Figure 3).

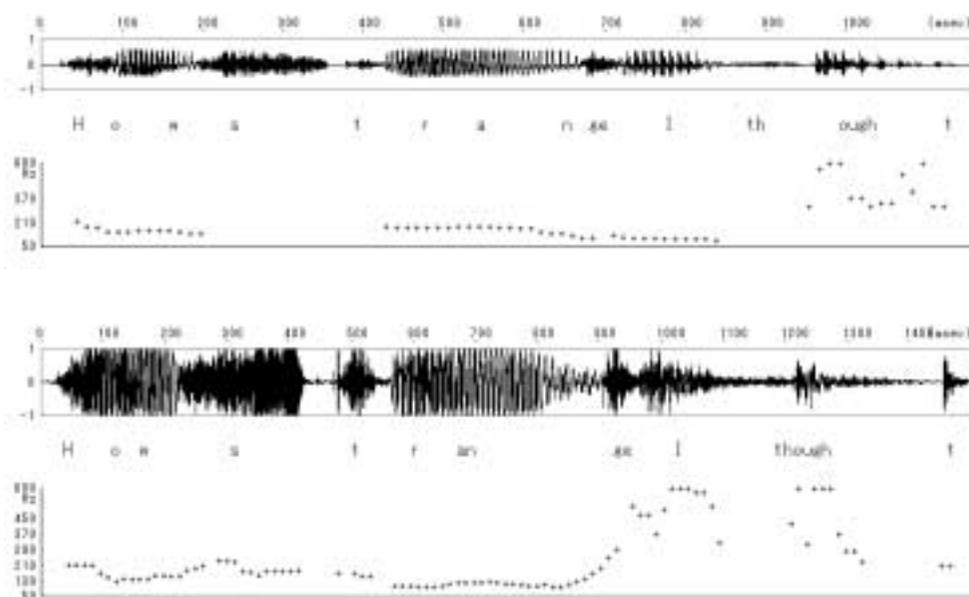


Figure 3. Above: Pitch contours of an American English model. Below: Pitch contours of British English model.

4. Discussion

As English pronunciation has various components such as individual sounds, stress, and variation in pitch, it is hard to judge whether a speaker's pronunciation has improved or not by measuring the duration of a consonant and comparing pitch contours. However, at a recent presentation (JALT 2005), my demonstration of random samples of the sound data of five students (AP and BP) got positive remarks from the floor. Of the 24 attendants at the session, 18 were native speakers of English and all of

them agreed that there was some improvement in pronunciation.

This study's objective is not to determine whether computer-assisted training or a particular software program is better than any other approach. The question might arise as to whether other non computer-based training approaches such as the traditional teacher-led instruction would be equally, or perhaps more, effective, and one might be inclined to attempt to compare these approaches experimentally. However, I would suggest that such a comparison is inherently difficult. As there are numerous elements that make up a training approach, all but the specific one under investigation would need to be the same in both approaches to avoid a confound. Simply using the same materials for the same period of time would not provide a suitable basis of comparison.

For some learners, technology holds greater interest than traditional approaches, which influences their motivation. Some enjoy a greater comfort level in working with a computer program than in face-to-face interaction where other personality factors are involved. While not an exhaustive list, the above points serve to emphasize that comparison of approaches, in general, is highly problematic. Note, however, that these comments are not intended as a claim that a particular type of training is best for all learners nor that instructor-led approaches are not beneficial, only that direct comparison is not well-founded.

As the training program was explicitly designed to deal with prosody first, students have not completed work on segmental features such as individual sounds. Further longitudinal investigation is needed to conclude whether or not prosodic cues facilitate the accuracy of pronouncing individual sounds. So far, results suggest that the training led to improved production at both the segmental and suprasegmental levels. Further experiments involving more learners at different levels of proficiency would contribute to our understanding of the potential of this approach. One learner commented that, "with practice I am getting more confident in speaking English. I felt insecure when I spoke English because sometimes people didn't understand what I said." The implication of this statement is that the fear of producing unintelligible speech makes learners' attempts at conversing frustrating and unpleasant. Pedagogically speaking, teaching English pronunciation must be based on the intelligibility and acceptability of learner's performance in a free discourse situation with a native or non-native speaker of the language. Frequent training opportunities should be provided to learners, and computer-assisted pronunciation training is one potentially valuable way of doing so.

Notes

- 1) Site name and URL: English Pronunciation Practice for Japanese Learners (<http://kccn.konan-u.ac.jp/ilc/english/>)
- 2) The short passage:
My mother once told me that she did not eat her first banana until she was twenty years old. How strange, I thought, especially when one considers the huge variety of fruits we can now find in a supermarket.

References

- 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.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge, England: Cambridge University Press.
- Chun, D. M. (1998). Signal analysis software for teaching discourse intonation. *Language Learning & Technology*, 2(1), 61-77. Retrieved July 22, 2002, from <http://lt.msu.edu/vol2num1/article4/>
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410.
- Di Cristo, A. (1998). Intonation in French. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 195-218). Cambridge, England: Cambridge University Press.
- Goldinger, S. D. (1997). Words and voices: Perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp. 33-66). San Diego, CA: Academic Press.
- Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context and talker variability. *Applied Psycholinguistics*, 24, 495-522.
- Hardison, D. M. (2000). The neurocognitive foundation of second-language speech: A proposed scenario of bimodal development. In B. Swierzbins, F. Morris, M. E. Anderson, C. A. Klee, & E. Tarone (Eds.), *Social and cognitive factors in second language acquisition* (pp. 312-325). Somerville, MA: Cascadilla Press.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
- Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review*, 93(4), 411-428.
- Hirst, D. J., & Di Cristo, A. (1998). A survey of intonation systems. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 1-44). Cambridge, England: Cambridge University Press.
- Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), *New Sounds 90* (pp. 72-97). Amsterdam: University of Amsterdam.
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242-1255.

- Molholt, G. (1988). Computer-assisted instruction in pronunciation for Chinese speakers of American English. *TESOL Quarterly*, 22(1), 91-111.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481-520.
- Munro, M. J. (1995). Nonsegmental factors in foreign accent. *Studies in Second Language Acquisition*, 17(1), 17-33.
- Pennington, M. C., & Esling, J. H. (1996). Computer-assisted development of spoken language skills. In M. C. Pennington (Ed.), *The power of CALL* (pp. 153-189). Houston, TX: Athelstan.
- Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL Quarterly*, 20, 207-225.
- Weltens B., & de Bot, K. (1984). Visual feedback of intonation II: Feedback delay and quality of feedback. *Language and Speech*, 27(1), 79-88.