Teaching and learning English prosody: Lessons from L2 speech perception and production research

CHAPTER · JANUARY 2012

Angelos Lengeris
Aristotle University of Thess…

Chapter 3
Prosody and Second Language Teaching: Lessons from L2 Speech Perception and Production Research

Angelos Lengeris

3.1 Introduction

Pronunciation accuracy in a second language (L2) requires mastering production of both segmental (i.e., consonants and vowels) and suprasegmental or prosodic features of speech (i.e., features that extend over more than one segment such as lexical stress, pitch accent, rhythm and intonation) but teaching pronunciation of the latter is traditionally neglected in language classrooms. After the advent of the communicative approach to language teaching (e.g., Celce-Murcia et al. 1996; Morley 1991, 1994) that prioritized language function over language form, the study of L2 prosody has admittedly experienced an increasing interest among language teachers. In addition, following Pierrehumbert's (1980) pioneer work, research in intonation is one of the most fast-growing areas in linguistics with the autosegmental-metrical theory being the dominant framework in intonational research. Studies comparing the relative contribution of segmental *vs*. prosodic features of speech in degree of foreign accent have shown that deviations in the latter may affect listeners' judgement more than deviations in the former (e.g., Munro 1995; Munro and Derwing 1999; Derwing et al. 1998). Specifically, prosody has been found to be linked to accentedness, comprehensibility and intelligibility of speech (Anderson-Hsieh et al. 1992; Anderson-Hsieh and Venkatagiri 1994; Hahn 2004; Jilka 2000; Kang 2010; Kang et al. 2010; Munro and Derwing 2001; Pickering 2001; Trofimovich and Baker 2006). These findings are not surprising considering that prosody and intonation in particular plays a crucial role in communication by conveying not only linguistic information such as chunking the stream of speech in phrases, signalling new and contrastive information and disambiguating sentences that otherwise could sound ambiguous to the listener, but also paralinguistic information, i.e., information related to the

A. Lengeris (✉)
Department of Theoretical and Applied Linguistics, School of English, Aristotle University of Thessaloniki, Thessaloniki, Greece
e-mail: lengeris@enl.auth.gr

J. Romero-Trillo (ed.), *Pragmatics and Prosody in English Language Teaching*, Educational Linguistics 15, DOI 10.1007/978-94-007-3883-6_3,

identity, age, gender, and emotional state of the speaker. Misunderstandings due to the use of wrong intonation may even lead to negative evaluation and discrimination (e.g., Munro 2003).

Despite all this work showcasing the importance of prosody in L2 learning, its teaching is commonly ignored in the curriculum. A still popular view among teachers and learners holds that pronunciation and above all intonation cannot be taught, especially after the learner has passed what is considered to be the critical period for language acquisition. In addition, the majority of language teachers are non-native speakers of the target language and may lack the confidence or the ability to reproduce the prosodic patterns in a native-like manner. At the same time, L2 speech perception and production findings are usually disseminated only to academic audiences and do not reach the classroom. Even if they did, researchers and language practitioners do not necessarily share the same interests nor are research findings always presented in such a way as to facilitate implementation in the teaching curriculum. This article attempts to address these issues by reviewing important findings from L2 speech perception and production research indicating that (a) L2 learning difficulties are caused by native language (L1) experience and not because of a maturational-based loss in neural plasticity, which leaves the window for learning open well into adulthood and (b) the human brain can be re-trained to perceive and produce L2 segmentals and suprasegmentals using appropriate computer-based techniques developed and tested for their effectiveness in the laboratory over the last two decades.

3.2 Age and Second Language Learning

Young learners are better in acquiring an L2 than older learners. It is a common belief among teachers, policy makers and researchers since Lenneberg (1967) published the *Biological Foundations of Language* introducing the concept of a critical period for language acquisition that this decline in L2 performance is due to an age-related change in neural plasticity. It has therefore been claimed that biologically determined maturational constraints exist when learning the L2 grammar (Johnson and Newport 1989), syntax (Patkowski 1980) and pronunciation (Patkowski 1990). Age effects on L2 learning are reported in numerous studies examining the perception and production of vowels (e.g., Flege et al. 1999a) and consonants (e.g., Mackay et al. 2001). Studies concerned with the effect of age on the learning of L2 suprasegmentals are more limited compared to the segmental ones. The majority of these studies focus on degree of global foreign accent, a measure that combines segmental and suprasegmental aspects of speech (e.g., Flege et al. 1999b; Oyama 1976) confirming a decline in learners' performance with age; old learners are found to have stronger foreign accents than early learners. In a recent study, Huang and Jun (2009) examined the age effect on the acquisition of various aspects of American English prosody by Chinese learners. Three groups of Chinese learners participated, varying in their age of arrival in the United States while the length of residence in the United States did not differ among groups. The study investigated Chinese learners' rate of English speech, the degree of foreign accent when speaking English (using low-pass filtered speech to remove segmental information while preserving the prosodic information) and the intonation patterns and prosodic groupings of their English speech production. The results confirmed an age effect on the acquisition of English prosody although the magnitude of the effect varied among the aspects of prosody that were examined.

The above studies demonstrate indisputable age effects on the acquisition of L2 segmentals and suprasegmentals. However, to support the view that such effects are due to an age-related loss in neural plasticity as proposed by the critical-period hypothesis, evidence is needed that (a) there is a sharp drop-off in the ability to learn a second language; (b) all early L2 learners achieve native-like performance; and (c) all late L2 learners fail to achieve native-like performance. On the contrary, a number of studies have shown that the perceptual system remains plastic enough to support learning well into adulthood and that there is no sharp drop-off in L2 learning ability but rather a gradual decline with age (Flege et al. 1999a, b). For example, Flege et al. (1999b) found that native Korean immigrants' degree of foreign accent increased as their age of arrival in the United States increased but there was no evidence of nonlinearity in Korean immigrants' performance. Further, other studies report that not all early bilinguals perform equally well (Flege et al. 1995, 1997) and that late bilinguals may achieve native-like pronunciation (e.g., Bongaerts et al. 1995, 1997; Moyer 1999). For example, Bongaerts et al. (1995, 1997) tested highly motivated Dutch learners of British English. None of the participants had received formal instruction in English before the age of around 12 and they were all exposed to a large amount of authentic L2 input delivered by native English speakers after entering the university. In the first study, foreign accent ratings were obtained for spontaneous speech, a text, 10 sentences and 25 words while in the second study ratings were obtained for 6 sentences. Bongaerts et al. (1995) found that all 10 Dutch participants were indistinguishable from native English speakers. Similarly, Bongaerts et al. (1997) found that 5 out of 11 participants met a criterion of 'native-likeness', i.e., their English sentence production received a mean rating that fell within 2 standard deviations of the mean rating obtained by English native speakers that were used as controls.

3.3 Linguistic Experience and Second Language Learning

If age-related changes in neural plasticity are not responsible for difficulties in learning the L2 segmentals and suprasegmentals, then what is it that makes it such a challenging task? Researchers believe that the advantage of early over late L2 learners is caused by our experience with our native language; as we grow up and acquire the sound system of our native language, our ability to learn patterns that are different from the native ones inevitably declines. A change in infants' perceptual abilities during the first year of life has been extensively described in a number of studies conducted the past 30 years. It has therefore been shown that, in the early months of life, infants are able to discriminate all sounds that are used to signal contrasts in any language (Aslin et al. 1981; Eimas et al. 1971; Trehub 1976). However, by the end of their first year infants fail to discriminate non-native consonant contrasts (Werker et al. 1981; Werker and Tees 1983, 1984). Sensitivity to non-native vowel contrasts appears to decline somewhat earlier, at around 6 months of age (Kuhl et al. 1992; Polka and Werker 1994). During their first year of life, infants demonstrate a similar perceptual reorganization for suprasegmental features of speech such as rhythm (Jusczyk et al. 1993) and lexical tone (Mattock and Burnham 2006, Mattock et al. 2008). For example, Jusczyk et al. (1993) showed that 9-month-old American infants, in contrast to 6-month-old American infants prefer to listen to words with a strong/weak stress pattern, which is the most frequently used pattern in English over words with a weak/strong pattern, indicating that experience with the prosodic features of the ambient language affects infants' response to language.

The role of L1 'tuning' in L2 speech learning is discussed in current L2/cross-linguistic models, such as the Perceptual Assimilation Model (Best 1995; Best and Tyler 2007), the Speech Learning Model (Flege 1995, 2002), and the Native Language Magnet model (Kuhl et al. 1992, 2008; Kuhl 2000). All three models agree that L1 language experience interferes with L2 learning and that the relationship between the L1 and L2 sound inventories can predict whether or not a specific L2 sound (or a specific L2 contrast) will pose difficulty to the learner. For example, the Speech Learning Model posits that as the L1 categories develop with age they become more powerful attractors of the L2 categories (e.g., Walley and Flege 1999). Several studies have demonstrated the role of linguistic experience in learning the L2 vowels (e.g., Cebrian 2006; Flege et al. 1999a; Flege and MacKay 2004; Iverson and Evans 2007; Lengeris 2009; Polka 1995) and consonants (e.g., Best et al. 2001; Guion et al. 2000; Hattori and Iverson 2009; Iverson et al. 2003; Mackay et al. 2001). For example, Spanish and Greek learners of English show a very poor discrimination of the English tense-lax vowel contrast /iː/-/ɪ/ because they lack such a contrast in their L1, having a single vowel category in the F1/F2 vowel space occupied by the two English vowels (Cebrian 2006 for Spanish learners; Lengeris 2009 for Greek learners). Likewise, Japanese speakers are very poor at differentiating English /r/ from /l/ because they pay attention to the non-critical second formant frequency (which is important for the perception of the Japanese voiced tap /ɾ/) instead of the critical third formant frequency (Iverson et al. 2003).

Effects of linguistic experience are also reported in studies on suprasegmental features of speech, specifically on the acquisition of stress (e.g., Archibald 1993; Dupoux et al. 1997, 2001; Guion et al. 2004; Peperkamp and Dupoux 2002; Peperkamp et al. 2010; Yu and Andruski 2010) and tone (e.g., Gottfried and Suiter 1997; Hallé et al. 2004; So and Best 2010; Wayland and Guion 2004). Clear effects of L1 experience on the way learners perceive and produce the L2 intonational patterns are also reported. Early studies focused on the errors produced by learners (e.g., Backman 1979; Willems 1982) but contemporary research has acknowledged the need to adopt a generally agreed framework for intonational analysis to better examine cross-linguistic similarities and differences in intonation. Mennen (2006) discusses the difficulties in comparing the findings of different studies in intonation research, as well as the potential of the Autosegmental framework of intonational analysis (Pierrehumbert 1980) in investigating L2 intonation. The model distinguishes between the underlying phonological representation of intonation (e.g., tonal inventory) and its phonetic manifestation (e.g., F0 peak alignment), providing a test-bed for the acquisition of L2 intonational targets and their phonetic realization. It is therefore not surprising that a growing number of studies have begun using the Autosegmental framework to examine the influence of L1 on the learning of L2 intonation during the last decade (e.g., Atterer and Ladd 2004; Jilka 2000; Mennen 2004, 2006).

Mennen (2004) investigated the extent to which the L1 intonation system can exert an influence on the acquisition of L2 intonation at the phonetic level. The study examined the production of Greek pre-nuclear rises by advanced Dutch learners of Greek. Greek and Dutch use phonologically identical pre-nuclear rises in declarative sentences but there are cross-linguistic differences in the phonetic manifestation of the rise. In Greek, the alignment of the peak is realized in the vowel following the accented syllable, whereas in Dutch the peak is realized slightly earlier, within the accented syllable. Furthermore, in Dutch the alignment of the peak is affected by the length of the vowel of the accented syllable (i.e., earlier when the vowel is long and later when the vowel is short), whereas in Greek it is not (there are no short-long distinctions in Greek). Five Dutch learners of Greek, all teachers of Greek at University level, participated in the study. The production of pre-nuclear rises by a group of Dutch native speakers and a group of Greek native speakers with no knowledge of Greek and Dutch respectively, were recorded for control reasons. Mennen (2004) found that four out of five Dutch learners of Greek transferred their L1 (Dutch) phonetic realization of pre-nuclear rises when speaking Greek (i.e., they aligned the peak earlier than Greek speakers) and only one Dutch learner managed to show native-like performance. Interestingly, Mennen (2004) reports a bi-directional interference in the production of pre-nuclear rises by those four Dutch learners of Greek; not only did they differ from Greek controls in their production of L2 (Greek) intonation but they also differed from Dutch controls in their production of L1 (Dutch) intonation. Only one Dutch learner managed to achieve native-like performance in peak alignment in both languages.

3.4 Learning in Naturalistic and Formal Settings

The role of L2 experience – usually indexed by the length of residence (LOR) in an L2 setting – in the acquisition of the L2 segmentals and suprasegmentals has been extensively examined in the literature but there are inconsistencies in findings across studies (see Piske et al. 2001 for a review of factors affecting degree of foreign accent in an L2). Some studies have found evidence supporting the importance of experience on L2 learning (e.g., Asher and Garcia 1969; Flege and Fletcher 1992;

Flege et al. 1997) while others report no such effect (e.g., Cebrian 2006; Oyama 1976; Piper and Cansin 1988). Trofimovich and Baker (2006) studied the effect of experience on the production of five English suprasegmentals (stress timing, peak alignment, speech rate, pause frequency and pause duration) by three groups of Korean learners of English who had been immersed in the United States after puberty and differed in their length of immersion (3 months, 3 and 10 years). Participants performed a delayed repetition task (declaratives as responses to question prompts). The sentences produced by Korean speakers were low-pass filtered to remove segmental effects and were rated by native English speakers for degree of global foreign accent. The sentences were also measured acoustically in terms of the five target suprasegmentals. All Korean speakers were found to be more accented than a control group of English speakers but those who were less experienced (i.e., 3 months of residence) were more accented than those with 3 and 10 years of residency (but there was no difference between the last two groups). The acoustic analysis showed that the amount of L2 experience correlated to stress timing but not to the other four suprasegmentals tested. In a following study, Trofimovich and Baker (2007) studied the effect of experience on the production of the same five English suprasegmentals by Korean learners of English who had been immersed in the United States before puberty and differed in their L2 experience (1 *vs*. 11 years of residency). The results showed that the latter group outperformed the former in all 5 suprasegmentals as well as in degree of global foreign accent. Those Koreans with 11 years of residency achieved native-like levels of performance in the global foreign accent task and in four suprasegmentals (all except speech rate).

Flege and Liu (2001) suggested that the lack of an effect of LOR in some studies may have been due to the quality of the L2 input the sampling population received. In their study, Flege and Liu (2001) examined the effect of LOR on L2 learning by means of a consonant identification task, a grammaticality judgment task and a listening comprehension task. The participants were adult Chinese speakers who had lived in the United States from 0.5 to 3.8 years (short LOR group) and from 3.9 to 15 years (long LOR group). Half of the participants in each group were university students while the remaining participants had worked full-time during their stay in the US. In all three tasks, an effect of LOR was found for the group of students but not for the non-students; only the former group achieved higher scores following immersion, a finding which demonstrates that L2 learning depends on the quality of native-speaker input that the learner receives (the two groups did not differ in terms of self-reported percentage use of English). Flege (2009) further discussed the importance of input in L2 learning. According to the author, both quality and quantity of input are important; residence in a foreign country is likely to be beneficial only for immigrants who receive a sufficient amount of L2 input via interaction with native speakers, especially via participation in social activities. In cases where immigrants receive a greater amount of L1-accented input than authentic input, the amount of L2 experience cannot be a reliable predictor of success in L2 learning.

Indirect evidence for the importance of authentic input when learning a second language comes from research in formal language settings (Elliott 1995; Fullana and MacKay 2010; Gallardo del Puerto et al. 2005; García Lecumberri and

Gallardo del Puerto 2003; MacKay and Fullana 2007; Mora and Fullana 2007). These studies report no effect of language instruction on L2 perception and production, which indicates that L2 exposure at a young age and several years of formal instruction may not lead to better pronunciation (see Singleton and Ryan 2004 for a review). This can be explained by the fact that classroom instruction is normally limited to a few hours per week, focuses on L2 form and the teacher who acts as a model to students delivers, in most cases, L1-accented input.

3.5 Learning in Laboratory Settings

One of the strongest arguments against the view that there is an age-related loss in neural plasticity comes from a number of computer-based training studies conducted over the past years. These studies have consistently shown that adults from various language backgrounds can be retrained to hear and produce L2 segmentals and suprasegmentals using structured intensive training procedures. Early studies attempting to modify perception of sounds adopted discrimination training whereby the trainees hear two synthetic stimuli in each trial and are asked to decide whether the two stimuli are identical or different (e.g., Carney et al. 1977; Pisoni et al. 1982; Strange and Dittmann 1984). These early studies showed a learning effect on the training stimuli but no transfer of learning to natural tokens. This was attributed partly to the use of discrimination training and partly to the low variability of the training stimuli. Regarding the former issue, it has been claimed that discrimination training tends to tailor learners' attention to within-category differences instead of focusing on between-category differences that are crucial for categorization. Regarding the latter issue, it is believed that the use of a single talker and context impedes transfer of learning to other talkers and contexts.

An alternative approach to training that has dominated the field over the past 20 years is the high-variability phonetic training technique. As its name indicates, this method emphasizes the importance of exposure to natural minimal pairs contrasting the target sounds in multiple environments spoken by multiple talkers thus resembling real-world communication with native speakers of the target language. High-variability training consists of an identification task with feedback whereby the trainees hear a single stimulus in each trial and are asked to label the sound using a number of given L2 categories. This approach to training has been found to significantly improve by about 20% the perception of L2 consonants and vowels (e.g., Logan et al. 1991; Lively et al. 1994; Hazan et al. 2005; Iverson and Evans 2009; Lengeris and Hazan 2010; Nishi and Kewley-Port 2007, 2008). Importantly, perceptual improvement is not limited to the stimuli used in training but generalizes to new words containing the target sounds and to talkers that were not heard during training. Perceptual learning is retained several months after training (Lively et al. 1994; Bradlow et al. 1999) and transfers to speech-in-noise vowel perception (Lengeris and Hazan 2010) and to consonant (Bradlow et al. 1997; Hazan et al. 2005) and vowel production (Lambacher et al. 2005; Lengeris and Hazan 2010).

Among the L2 segmentals that have been trained are the English vowel distinctions for Spanish, Greek, German and French speakers, the English /r/-/l/ distinction for Japanese speakers, the English word final /t/-/d/ distinction for Chinese speakers and the Hindi dental-retroflex stop distinction for English and Japanese speakers. Note that consonant studies usually train binary L2 contrasts (e.g., /r/-/l/, or /t/-/d/) whereas vowel studies have successfully trained several L2 vowels at the same time.

Wang and Munro (2004) examined whether computer-based training procedures such as these reviewed so far can be effective in improving L2 vowel perception in a classroom setting whereby learners are given some control over training. The authors trained Mandarin and Cantonese speakers on three English vowel contrasts /iː/-/ɪ/, /e/-/æ/, and /ʊ/-/uː/ using synthetic and natural vowel stimuli. The trainees were asked to schedule their 50–60 min training sessions (2–3 per week) and in each session they could recycle the training blocks as desired. After training, participants improved their perception of English vowels; their improvement was transferred to a new context and was retained 3 months after training. The results of this study are important because they show that research knowledge from laboratory training studies can be put into practice in a classroom environment.

Another line of research has been exploring the implementation of fundamental frequency visualization software for teaching L2 intonation since the 1970s (Abberton and Fourcin 1975; Anderson-Hsieh 1992, 1994; Chun 1998; De Bot 1983; Hardison 2004; Levis and Pickering 2004; Spaai and Hermes 1993; Taniguchi and Abberton 1999; Weltens and de Bot 1984). Early software (and hardware) used to be expensive and difficult to use but nowadays computer technology for speech analysis is becoming widely accessible. A number of analysis programs developed by the research community such as Praat (Boersma and Weenink 2009), SFS (Huckvale 2008), Wavesurfer (Sjölander and Beskow 2000) and Speech Analyzer (SIL 2007) are freely available online. Learners can record themselves and see on the screen a visual representation of the pitch contour of their speech. The visualization of pitch is relatively easy to interpret by inexperienced learners and does not require extensive phonetic knowledge, as is required for the interpretation of e.g., spectrograms or waveforms. Learners can also compare their production both auditory and visually with a model speaker, which helps to raise learners' awareness of differences between L1 and L2 intonation patterns.

Anderson-Hsieh (1992) trained international teaching assistants (mainly Chinese and Korean native speakers) on English word stress, rhythm, linking and intonation using visual feedback. Although the paper's aim was to demonstrate how visual feedback can be used as a tool for teaching suprasegmentals rather than evaluating the effectiveness of training using statistical analyses, the author reports that training had a positive effect on the learners' performance and attributes this finding to the fact that a visual representation of suprasegmentals in real time accompanied the auditory feedback provided to the learners. De Bot (1983) compared the effects of audio-visual feedback and audio-only feedback on Dutch students' learning of English intonation. The results showed that the group of students who received audio-visual feedback improved more than those who received audio-only feedback, as judged by three teachers of English. Taniguchi and Abberton (1999) examined the effectiveness of interactive visual feedback on Japanese speakers' production of English intonation. The Japanese speakers were 12 college students who attended the UCL Summer Course in English Phonetics. All students attended lectures and practical lessons in English intonation but only half of them received training with visual feedback during their intonation lessons. All 12 Japanese speakers improved their English intonation after 2 weeks but those who received interactive visual feedback outperformed those who did not receive visual feedback.

Hardison (2004) examined the effectiveness of computer-assisted training on the acquisition of French prosody by native speakers of American English using pitch displays and multiple recordings spoken by multiple native French speakers as feedback. This training protocol differs from protocols used in suprasegmental training studies reviewed so far, as native speakers of the target language were used as feedback and not as models to imitate. Hardison's study consisted of two experiments. In Experiment 1, 26 American English speakers, all undergraduate university students, participated. Of those speakers, 16 were trained while the remaining speakers served as controls, i.e., performed only the pre/post test without receiving any training. In the pre/post test, participants were asked to produce 20 French sentences while another 20 novel sentences in the post-test evaluated generalization of learning. Training consisted of thirteen 40-min sessions with sets of 30 sentences spoken by three native French speakers. During each training session, participants were asked to produce one set of 30 sentences. The pitch contours of their sentences were displayed in real time on the screen. A French speaker's version of that sentence was displayed on the screen in a different colour and played out aloud. Three native French speakers evaluated the recordings made by the American English speakers on a 7-point scale. The results indicated that the trainees improved their production and that this improvement generalized to the set of novel sentences. Experiment 2 used a memory recall task using filtered stimuli that preserved prosodic information while reducing segmental information to examine whether learners were able to identify the exact lexical content of the training sentences based on prosodic information alone. The trainees succeeded about 80% of the time in doing so, demonstrating according to the author that prosodic and lexical information is stored together in memory traces. Finally, participants' responses from questionnaires indicated a greater awareness of the various aspects of speech after training and increased confidence when speaking French.

Following the growing interest in discourse intonation among intonation experts (e.g., Bolinger 1989; Brazil 1997; Chun 2002) and the success of intonation training studies that have used sentence-level training materials, researchers have begun to explore ways to teach intonation patterns that occur in communicative contexts (Chun 1998; Levis and Pickering 2004). For example, Hardison (2005) examined whether computer-based training can improve the production of L2 prosody at the level of contextualized speech, using pitch displays and discourse-level training materials. Twenty-eight Chinese advanced learners of English participated in the study. Half of the participants were trained on discourse-level materials and half were trained on individual sentences. Three native English speakers provided global prosody ratings for both groups of Chinese speakers. The results showed that all speakers showed transfer of learning to natural discourse as a result of training and that the group of speakers that received contextualized input improved more than the group that received sentence-level input. Prosody ratings of speech materials produced by the same Chinese learners 1 week after training revealed sustained improvement. These results showcase that computer technology can be effective in teaching not only the typical sentence-type intonation patterns (e.g., declaratives, wh-questions, yes-no questions, etc.) but also discourse-level intonation patterns.

3.6 Conclusion

The acquisition of second language prosody is undeniably a difficult task. However, a growing body of experimental work supports the view that, as with segmental aspects of speech, the window of learning is not closed even in adulthood. This is because difficulties in L2 learning are mainly caused by native language experience and not because of an age-related change in neural plasticity that would make learning unfeasible. Strong evidence for this plasticity comes from laboratory studies showing that the adult brain can be trained to hear non-native differences by using appropriate computer-based training techniques. What is even more encouraging is that perceptual training can improve not only the perception of L2 segmentals and suprasegmentals but also their production. The most successful training protocols adopt the so-called high-variability approach, which has been proven to create robust and long-lasting new categories. In line with exemplar and statistical models of speech perception (Goldinger 1996; Johnson 1997; Pierrehumbert 2002; Maye et al. 2002), in order to create new categories we simply need to expose the learner to multiple natural tokens of the target sounds produced by various speakers. Exposure to authentic and variable L1 input is therefore vital in the learning process. In the case of suprasegmentals, the visualization of the pitch contour can help those learners who cannot rely only on their own perception of intonation to improve their production of both sentence- and discourse-level prosody. The area of L2 intonation training thus provides a perfect example of how valuable insights from academic work can be put into practice to assist the teaching of one of the most difficult aspects of second language pedagogy.

References

Abberton, E., and A. Fourcin. 1975. Visual feedback and the acquisition of intonation. In *Foundations of language development*, ed. E.H. Lenneberg and E. Lenneberg, 157–165. New York: Academic Press.

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

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

Anderson-Hsieh, J., and H. Venkatagiri. 1994. Syllable duration and pausing in the speech of intermediate and high proficiency Chinese ESL speakers. *TESOL Quarterly* 28: 807–812.

Anderson-Hsieh, J., R. Johnson, and K. Koehler. 1992. The relationship between native speaker judgments of non-native pronunciation and deviance in segmentals, prosody and syllable structure. *Language Learning* 42: 529–555.

Archibald, J. 1993. The learnability of English metrical parameters by adult Spanish speakers. *International Review of Applied Linguistics* 31: 129–141.

Asher, J., and R. Garcia. 1969. The optimal age to learn a foreign language. *The Modern Language Journal* 53: 334–341.

Aslin, R.N., D.B. Pisoni, B.L. Hennessy, and A.J. Perey. 1981. Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development* 52: 1135–1145.

Atterer, M., and D.R. Ladd. 2004. On the phonetics and phonology of "segmental anchoring" of F0: Evidence from German. *Journal of Phonetics* 32: 177–197.

Backman, N.E. 1979. Intonation errors in second language pronunciation of eight Spanish speaking adults learning English. *Interlanguage Studies Bulletin* 4(2): 239–266.

Best, C.T. 1995. A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. W. Strange, 171–204. Baltimore: York Press.

Best, C.T., and M.D. Tyler. 2007. Non-native and second-language speech perception: Commonalities and complementarities. In *Second language speech learning*, ed. M. Munro and O.-S. Bohn, 13–34. Amsterdam: John Benjamins.

Best, C.T., G.W. McRoberts, and E. Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109: 775–794.

Boersma, P., and D. Weenink. 2009. Praat: Doing phonetics by computer. [Computer Program]. http://www.praat.org/. Accessed 12 Aug 2009.

Bolinger, D. 1989. *Intonation and its uses: Melody in grammar and discourse*. London: Arnold.

Bongaerts, T., B. Planken, and E. Schils. 1995. Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In *The age factor in second language acquisition*, ed. D. Singleton and Z. Lengyel, 30–50. Clevedon: Multilingual Matters.

Bongaerts, T., C. Van Summeren, B. Planken, and E. Schils. 1997. Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition* 19: 447–465.

Bradlow, A.R., D.B. Pisoni, R. Akahane-Yamada, and Y. Tohkura. 1997. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101: 2299–2310.

Bradlow, A.R., R. Akahane-Yamada, D.B. Pisoni, and Y. Tohkura. 1999. Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics* 61: 977–985.

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

Carney, A.E., G.P. Widin, and N.F. Viemeister. 1977. Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America* 62: 961–970.

Cebrian, J. 2006. Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics* 34: 372–387.

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

Chun, D.M. 1998. Signal analysis software for teaching discourse intonation. *Language Learning and Technology* 2: 61–77. http://llt.msu.edu/vol2num1/article4. Accessed 11 Mar 2009.

Chun, D.M. 2002. *Discourse intonation in L2: From theory and research to practice*. Amsterdam: John Benjamins.

De Bot, K. 1983. Visual feedback on intonation I: Effectiveness and induced practice behaviour. *Language and Speech* 26: 331–350.

Derwing, T.M., M.J. Munro, and G. Wiebe. 1998. Evidence in favour of a broad framework for pronunciation instruction. *Language Learning* 48: 393–410.

Dupoux, E., C. Pallier, N. Sebastian, and J. Mehler. 1997. A destressing “deafness” in French? *Journal of Memory and Language* 36: 406–421.

Dupoux, E., S. Peperkamp, and N. Sebastian. 2001. A robust method to study stress “deafness”. *Journal of the Acoustical Society of America* 110: 1606–1618.

Eimas, P.D., E.R. Siqueland, P. Jusczyk, and J. Vigorito. 1971. Speech perception in infants. *Science* 171: 303–306.

Elliott, A.R. 1995. Foreign language phonology: Field independence, attitude and the success of formal instruction in Spanish pronunciation. *The Modern Language Journal* 79: 530–542.

Flege, J.E. 1995. Second language speech learning: Theory, findings, and problems. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. W. Strange, 233–277. Baltimore: York Press.

Flege, J.E. 2002. Interactions between the native and second-language phonetic systems. In *An integrated view of language development: Papers in honor of Henning Wode*, ed. P. Burmeister, T. Piske, and A. Rohde, 217–244. Trier: Wissenschaftlicher Verlag.

Flege, J.E. 2009. Give input a chance! In *Input matters in SLA*, ed. T. Piske and M. Young-Scholten, 175–190. Bristol: Multilingual Matters.

Flege, J.E., and K.L. Fletcher. 1992. Talker and listener effects on the degree of perceived foreign accent. *Journal of the Acoustical Society of America* 91(1): 370–389.

Flege, J.E., and S. Liu. 2001. The effect of experience on adults’ acquisition of a second language. *Studies in Second Language Acquisition* 23: 527–552.

Flege, J.E., and I.R.A. Mackay. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26: 1–34.

Flege, J.E., M. Munro, and I.R.A. MacKay. 1995. Effects of age of second language learning on the production of English consonants. *Speech Communication* 16: 1–26.

Flege, J.E., E.M. Frieda, and T. Nozawa. 1997. Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics* 25: 169–186.

Flege, J.E., I.R.A. MacKay, and D. Meador. 1999a. Native Italian speakers’ production and perception of English vowels. *Journal of the Acoustical Society of America* 106: 2973–2987.

Flege, J.E., G.H. Yeni-Komshian, and S. Liu. 1999b. Age constraints on second language acquisition. *Journal of Memory and Language* 41: 78–104.

Fullana, N., and I.R.A. MacKay. 2010. L2 fluency in the production of English extemporaneous speech by Catalan/Spanish bilinguals. In *Proceedings of the Sixth International Symposium on the Acquisition of Second Language Speech*, 125–130, Poznan, Poland. http://ifa.amu.edu.pl/newsounds/. Accessed 5 Jan 2011.

Gallardo del Puerto, F., M.L. García Lecumberri, and J. Cenoz. 2005. Degree of foreign accent and age of onset in formal school instruction. http://www.phon.ucl.ac.uk/home/johnm/ptlc2005/pdf/ptlcp48.pdf. Accessed 2 Feb 2009.

García Lecumberri, M.L., and F. Gallardo del Puerto. 2003. English FL sounds in school learners of different ages. In *Age and the acquisition of English as a foreign language*, ed. M.P. García Mayo and M.L. García Lecumberri, 115–135. Clevedon: Multilingual Matters.

Goldinger, S.D. 1996. Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 1166–1183.

Gottfried, T., and T. Suiter. 1997. Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics* 25: 207–231.

Guion, S.G., J.E. Flege, R. Akahane-Yamada, and J.C. Pruitt. 2000. An investigation of current models of second language speech perception: The case of Japanese adults’ perception of English consonants. *Journal of the Acoustical Society of America* 107: 2711–2724.

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

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

Hallé, P., Y.-C. Chang, and C.T. Best. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics* 32: 395–421.
Hardison, D.M. 2004. Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning and Technology* 8: 34–52. http://llt.msu.edu/vol8num1/pdf/hardison.pdf. Accessed 22 Oct 2008).
Hardison, D.M. 2005. Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal* 22(2): 175–190.
Hattori, K., and P. Iverson. 2009. English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America* 125: 469–479.
Hazan, V., A. Sennema, M. Iba, and A. Faulkner. 2005. Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication* 47: 360–378.
Huang, B., and S. Jun. 2009. Age effect on the acquisition of second language prosody. In *Proceedings of the 33rd Boston University Conference on Language Development*. http://www.bu.edu/linguistics/APPLIED/BUCLD/supplement33/Huang.pdf. Accessed 6 Sept 2010.
Huckvale, M. 2008. Speech Filing System tools for speech research, University College London. [Computer program]. http://www.phon.ucl.ac.uk/resource/sfs/. Accessed 3 Apr 2009.
Iverson, P., and B.G. Evans. 2007. Learning English vowels with different first language vowel systems: Perception of formant targets, formant movement, and duration. *Journal of the Acoustical Society of America* 122: 2842–2854.
Iverson, P., and B.G. Evans. 2009. Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *Journal of the Acoustical Society of America* 126: 866–877.
Iverson, P., P.K. Kuhl, R. Akahane-Yamada, E. Diesch, Y. Tohkura, and A. Kettermann. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87: 47–57.
Jilka, M. 2000. The contribution of intonation to the perception of foreign accent. PhD thesis, University of Stuttgart, Stuttgart.
Johnson, K. 1997. Speech perception without speaker normalization: An exemplar model. In *Talker variability in speech processing*, ed. K. Johnson and J. Mullenix, 145–166. San Diego: Academic.
Johnson, J., and E. Newport. 1989. Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21: 60–99.
Jusczyk, P.W., A. Cutler, and N.J. Redanz. 1993. Infants' preference for the predominant stress patterns of English words. *Child Development* 64: 675–687.
Kang, O. 2010. Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System* 38: 301–315.
Kang, O., D. Rubin, and L. Pickering. 2010. Suprasegmental measures of accentedness and judgments of English language learner proficiency in oral English. *Modern Language Journal* 94: 554–566.
Kuhl, P.K. 2000. A new view of language acquisition. *National Academy of Sciences USA* 97: 11850–11857.
Kuhl, P.K., K.A. Williams, F. Lacerda, K.N. Stevens, and B. Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255: 606–608.
Kuhl, P.K., B.T. Conboy, S. Coffey-Corina, D. Padden, M. Rivera-Gaxiola, and T. Nelson. 2008. Phonetic learning as a pathway to language: New data and native language magnet theory expanded NLM-e. *Philosophical Transactions of the Royal Society* B363: 979–1000.
Lambacher, S.G., W.L. Martens, K. Kakehi, C.A. Marasinghe, and G. Molholt. 2005. The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied PsychoLinguistics* 26: 227–247.
Lengeris, A. 2009. Perceptual assimilation and L2 learning: Evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese. *Phonetica* 66: 169–187.

Lengeris, A., and V. Hazan. 2010. The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *Journal of the Acoustical Society of America* 128(6): 3757–3768.

Lenneberg, E.H. 1967. *Biological foundations of language*. New York: Wiley.

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

Lively, S.E., D.B. Pisoni, R. Yamada, Y. Tohkura, and T. Yamada. 1994. Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America* 96: 2076–2087.

Logan, J.S., S.E. Lively, and D.B. Pisoni. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89: 874–886.

Mackay, I.R.A, and N. Fullana. 2007. Starting age and exposure effects on EFL learners' sound production in a formal learning context. In *Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*, 324–335, Florianópolis, Brazil. http://www.nupffale.ufsc.br/newsounds/Papers/28.Mackay_Fullana.pdf. Accessed 11 May 2009.

Mackay, I.R.A., D. Meador, and J.E. Flege. 2001. The identification of English consonants by native speakers of Italian. *Phonetica* 58: 103–125.

Mattock, K., and D. Burnham. 2006. Chinese and English infants' tone perception: Evidence for perceptual reorganization. *Infancy* 103: 241–264.

Mattock, K., M. Molnar, L. Polka, and D. Burnham. 2008. The developmental course of lexical tone perception in the first year of life. *Cognition* 106(3): 1367–1381.

Maye, J., J.F. Werker, and L. Gerken. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101–B111.

Mennen, I. 2004. Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32: 543–563.

Mennen, I. 2006. Phonetic and phonological influences in non-native intonation: An overview for language teachers. *QMUC Speech Science Research Centre Working Papers* WP-9.

Mora, J.C., and Fullana, N. 2007. Production and perception of /i□/-/□/ and /æ/-/□/ in a formal setting: Investigating the effects of experience and starting age. In *Proceedings of the 16th International Congress of Phonetic Sciences*, 1613–1616, Saarbrücken, Germany.

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

Morley, J. 1994. A multidimensional curriculum design for speech/pronunciation instruction. In *Pronunciation theory and pedagogy: New views, new directions*, ed. J. Morley, 64–91. Alexandria: TESOL Publications.

Moyer, A. 1999. Ultimate attainment in L2 phonology. *Studies in Second Language Acquisition* 21: 81–108.

Munro, M.J. 1995. Non-segmental factors in foreign accent. *Studies in Second Language Acquisition* 17: 17–34.

Munro, M.J. 2003. A primer on accent discrimination in the Canadian context. *TESL Canada Journal* 202: 38–51.

Munro, M.J., and T. Derwing. 1999. Foreign accent, comprehensibility and intelligibility in the speech of second language learners. *Language Learning* 49(Supplement 1): 285–310.

Munro, M.J., and T.M. Derwing. 2001. Modelling perceptions of the comprehensibility and accentedness of L2 speech: The role of speaking rate. *Studies in Second Language Acquisition* 23: 451–468.

Nishi, K., and D. Kewley-Port. 2007. Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research* 50: 1496–1509.

Nishi, K., and D. Kewley-Port. 2008. Non-native speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research* 51: 1480–1493.

Oyama, S. 1976. A sensitive period for the acquisition of a non-native phonological system. *Journal of Psycholinguistic Research* 5: 261–283.

Patkowski, M.S. 1980. The sensitive period for the acquisition of syntax in a second language. *Language Learning* 30: 449–472.

Patkowski, M.S. 1990. Age and accent in second language: A reply to James Emil Flege. *Applied Linguistics* 11: 73–89.

Peperkamp, S., and E. Dupoux. 2002. A typological study of stress "deafness". In *Laboratory phonology 7*, ed. C. Gussenhoven and N. Warner, 203–240. Berlin: Mouton de Gruyter.

Peperkamp, S., I. Vendelin, and E. Dupoux. 2010. Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics* 38(3): 422–430.

Pickering, L. 2001. The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly* 35: 233–255.

Pierrehumbert, J. 1980. The phonology and phonetics of English intonation. PhD thesis, MIT, Indiana University Linguistics Club.

Pierrehumbert, J. 2002. Word-specific phonetics. In *Laboratory phonology VII*, ed. L. Phonology, 101–139. Berlin: Mouton de Gruyter.

Piper, T., and D. Cansin. 1988. Factors influencing the foreign accent. *The Canadian Modern Language Review* 44: 334–342.

Piske, T., I.R.A. Mackay, and J.E. Flege. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29: 191–215.

Pisoni, D.B., R.N. Aslin, A.J. Perey, and B.L. Hennessy. 1982. Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology. Human Perception and Performance* 8: 297–314.

Polka, L. 1995. Linguistic influences in adult perception of non-native vowel contrasts. *Journal of the Acoustical Society of America* 97: 1286–1296.

Polka, L., and J.F. Werker. 1994. Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology. Human Perception and Performance* 20: 421–435.

Singleton, D., and L. Ryan. 2004. *Language acquisition: The age factor*. Clevedon: Multilingual Matters.

Sjölander, K., and J. Beskow. 2000. WaveSurfer – An open source speech tool. In *Proceedings of the 6th International Conference on Spoken Language Processing*. http://www.speech.kth.se/wavesurfer/. Accessed 17 Dec 2009.

So, C.K., and C.T. Best. 2010. Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language and Speech* 53: 273–293.

Spaai, G.W.G., and D.J. Hermes. 1993. A visual display for the teaching of intonation. *CALICO Journal* 103: 19–30.

Speech Analyzer. SIL. 2007. [Computer program]. http://www.sil.org/computing/sa/index.htm. Accessed 4 Oct 2009.

Strange, W., and S. Dittmann. 1984. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics* 36: 131–145.

Taniguchi, M., and E. Abberton. 1999. Effect of interactive visual feedback on the improvement of English intonation of Japanese EFL learners. *Speech, Hearing and Language: Work in Progress* 11: 76–89.

Trehub, S.E. 1976. The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47: 466–472.

Trofimovich, P., and W. Baker. 2006. Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition* 28: 1–30.

Trofimovich, P., and W. Baker. 2007. Learning prosody and fluency characteristics of L2 speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied PsychoLinguistics* 28: 251–276.

Walley, A.C., and J.E. Flege. 1999. Effect of lexical status on children's and adults' perception of native and non-native sounds. *Journal of Phonetics* 27: 307–332.

Wang, X., and M.J. Munro. 2004. Computer-based training for learning English vowel contrasts. *System* 32: 539–552.

Wayland, R.P., and S.G. Guion. 2004. Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning* 54(4): 681–712.

Weltens, B., and K. de Bot. 1984. Visual feedback of intonation II: Feedback delay and quality of feedback. *Language and Speech* 271: 79–88.

Werker, J.F., and R.C. Tees. 1983. Developmental change across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology* 37: 278–286.

Werker, J.F., and R.C. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development* 7: 49–63.

Werker, J.F., J.H.V. Gilbert, K. Humphrey, and R.C. Tees. 1981. Developmental aspects of cross-language speech perception. *Child Development* 52: 349–355.

Willems, N.J. 1982. *English intonation from a Dutch point of view*. Dordrecht: Foris Publications.

Yu, V.Y., and J.E. Andruski. 2010. A cross-language study of perception of lexical stress in English. *Journal of Psycholinguistic Research* 39(4): 323–344.