Concordia Working Papers in Applied Linguistics

Proceedings of the International Symposium on the Acquisition of Second Language Speech
Concordia Working Papers in Applied Linguistics, 5, 2014 © 2014 COPAL

Mandarin Speakers' English L2 Prosody and Degree of Foreign Accent: Effect of Length of Residence

Xinchun Wang
California State University, Fresno

Abstract

This study examines how Mandarin speakers' L2 English speech prosody contributes to perceived degree of foreign accent. A secondary goal is to explore whether Mandarin speakers' English speech prosody is influenced by their extended length of residence in the U.S. The participants were two groups of highly advanced ESL speakers who used English on a daily basis in teaching. The two Mandarin groups were comparable in terms of learner characteristics such as age of learning, years of learning, and other L2 experience but differed significantly in Length of Residence (LOR) in the U.S. (0 vs. 12 years). The speakers' productions of English sentences and excerpts of spontaneous speech rated for degree of foreign accent by native English listeners were submitted for acoustic measurement of prosodic features. The instrumental analysis focused on speech rate, pause features and these temporal properties (independent variables) are compared with the results of foreign accent rating (dependent variable) to explore which prosodic features predict the degree of foreign accent. Results show that the two native Mandarin groups' speech rate did not differ from each other but was significantly slower than that of a native English control group in both spontaneous speech and read sentences. For spontaneous speech, pause features including number of pauses, mean pause duration, mean

percentage pause duration are the predictors of foreign accent when the speech rate factors were phrased out.

Adult second language (L2) speakers often speak the target language with a noticeable foreign accent. Research has shown that a number of learner factors, such as Age of Arrival (AOA), Length of Residence (LOR), the speakers' first language (L1), and L2 experience contribute to degree of foreign accent (Piske, MacKay, & Flege, 2001). While age of arrival has consistently been found to be the most important factor that influences the L2 speakers' perceived degree of accent (Flege, Munro, & MacKay, 1995; MacKay, Flege, & Imai, 2006), the effect of length of residence in the target language environment has been mixed. The mixed findings on the effect of LOR on degree of foreign accent might be due to the speakers' differences in different length of residence in different studies (Flege, 1988; Trofimovich & Baker, 2006) and different learning stages and proficiency levels the learners were at during the time of the study (Piske et al., 2001; Riney & Flege, 1998; Riney, Takagi, & Inutsuka, 2005; Trofimovich & Baker, 2007). In addition, learners' first language and the distance between L1 and L2 may also contribute to L2 speakers' perceived degree of accent (Bongaerts, van Summeren, Planken, & Schils, 1997, Piske et al., 2001).

Researchers working on L2 foreign accent have also investigated which speech properties affect the perceived degree of foreign accent. L2 speech temporal features such pauses, speech rate and other suprasegmental features including stress timing and pitch range, as well as segmental errors, have been found to be related to perceived degree of foreign accent (Kang, 2010, Munro, 1995; Trofimovich & Baker, 2006). The relative weights of individual temporal and prosodic features for listeners' judgments on L2 comprehensibility and accentedness have also been explored. In these studies, acoustic measurement of temporal and suprasegmental features were used as independent variables to predict the degree of comprehensibility and foreign accent (Kang, 2010; Kang, Rubin, & Pickering, 2010; Trofimovich & Baker, 2006). For example, in Kang's (2010) study, the international teaching assistants' lecture excerpts were rated for comprehensibility and foreign accent by native English listeners. The analyses of their speech prosodic features showed the speakers' foreign accent rating scores were best predicted by pitch range and word stress measures and their comprehensibility scores were mostly associated with speech rates. In another study, Trofimovich and Baker

(2006) examined the native Korean speakers' different L2 experience (three different groups with LOR of 3 months, 3 year, and 10 year) and their effect on their L2 English productions' perceived degree of foreign accent. Rating scores and speech property measurements were taken to determine the relative weights of each speech temporal and suprasegmental feature that contributed to the degree of accent. They found that suprasegmentals contributed to native Korean speakers' perceived degree of foreign accent at all levels of experience. In particular, pause duration and speech rate contributed more to foreign accent than stress timing and peak alignment. Age of arrival (AOA) influenced the speech rate, pause frequency, and pause duration. Munro and Derwing's study (2001) investigated the relationship between speech rate and degree of foreign accent by nonnative speakers with different L1 background. Using the read sentences that were manipulated with increased and decreased speed for foreign accent rating, they found the curvilinear effect of speaking rate on listeners' judgments of degree of foreign accent, which indicates that L2 speech that is either too fast or too slow could both negatively affect nonnative speakers' degree of accent. In fact they found that the optimal rate was somewhat faster than the rates typically used by the nonnative speakers. Other temporal features such as pauses and pause durations were not investigated.

Previous studies on Mandarin speakers' acoustic characteristics of L2 English productions mainly focused on lexical stress and contrastive stress patterns (Chen et al., 2001; Zhang, Nissan, & Francis, 2008). These studies explored whether native Mandarin speakers used the same acoustic cues such as pitch, intensity, and vowel reduction to signal English lexical stress the way native English speakers do and whether the differences in using different acoustic cues by native Mandarin speakers were attributed to their L1 experience. However, these studies did not directly measure how these speech property differences are related to Mandarin speakers' degree of foreign accent or comprehensibility. The relative weight of Mandarin speakers' English speech prosody, in particular, speech temporal properties that might contribute to perceived degree of foreign accent have not received much attention in previous studies.

The current study mainly investigates native Mandarin speakers' L2 English speech temporal features in read sentences and spontaneous speech and their relative weights on perceived degree of foreign accent. The other goal is to examine whether extended LOR in the target language environment has an effect on adult L2 speakers' speech prosody and their degree of foreign accent. The participants were two groups of advanced

Mandarin speakers that were controlled for other learning experience but differed in LOR. The research questions are:

1. Do Mandarin speakers produce different English prosodic temporal features such as speech rate and pauses as compared to native English speakers?
2. Which prosodic features predict Mandarin speakers' degree of foreign accent as judged by English listeners?
3. Does speakers' extended LOR make a difference in producing English prosody and does LOR affect their degree of foreign accent?

EXPERIMENT 1. FOREIGN ACCENT RATING

Methods

Speakers. Two groups of highly advanced native Mandarin learners of English participated as speakers.[1] They were the America Professor Group (the AP Group) and the Chinese Professor Group (the CP Group). The AP group consisted of 10 professors (6 male, 4 female, mean age = 43.4, range 33-52) teaching at a university in the U.S. at the time of this study. They were all born and raised in China and studied English as a foreign language for a mean of 10 years at school and earned at least one university degree in China before they moved to North America. Their mean age of arrival (AOA) in North America was 31.3 years (23-37). Their mean LOR in North America was 12.4 years (5-21). All had earned at least one university degree in North America. They had taught at college level for a mean of 12.4 years (2-18) in various subject areas including science, engineering, business, education, and humanities. As expected, the AP group reported a very high mean percentage use of English, 97.5% (90-100%), outside the home but a mean of 37% (5%-95%) use of English at home.

The CP Group consisted of 10 professors (4 male, 6 female, mean age = 40.3, range 33-59) in a university in China. All had at least one university degree in English or English linguistics from China but none had earned a degree from an English speaking country. Some reported having taken an English phonetics course when studying for their English or linguistic

[1] Part of this experiment was presented at ICPhS 2007 and published in the Proceedings of ICPhS.

majors. None had lived in an English speaking country for over 6 months and many had never been abroad by the time of this study. They taught English or English linguistics in a university in China for a mean of 13.2 years (4-25). Though all spoke English fluently, they reported seldom using English at home (mean use = 2.5%, range 0%-10%). However, they all spoke English exclusively in class while teaching and generally used English to prepare their English and linguistics courses. (See Wang & Chen, 2007 for AP and CP groups' background information.)

Five native English speaker professors (3 male, 2 female, mean age = 40) formed a control group, the EP Group. They were all from the same university where the AP group was recruited.

Listeners. The listeners were 11 native English speakers (4 male, 7 female, mean age = 27.6, range: 21-37) from the student population at the university where the AP and EP Group were recruited but none were familiar with any of the speakers. None reported being familiar with Mandarin accent. All reported having normal hearing.

Stimuli. The stimuli included sentences, spontaneous speech (utterances) and low pass filtered spontaneous speech (filtered utterances). The sentences were elicited through a reading task in which the speakers read a list of 10 short sentences at normal speed. Only two of the ten, one statement and one question, were used for rating: "most people like to listen to music", "do you have a gas cooker in your kitchen?" The spontaneous speech was elicited through interviews in which the speakers answered two related questions: "do you enjoy teaching?" and "what do you like the most about teaching?" The speakers heard the questions from the researcher twice and were given 1-2 minutes to prepare before they provided the answers. They were told to use the questions as the prompt for their talk. The recordings were made in a quiet room using a Sanyo mini cassette recorder (TRC-680MN) and an external microphone. The recorded sentences and interviews were digitized at a sampling rate of 22050 Hz with 16-bit resolution and normalized for peak intensity. For the spontaneous speech, a short excerpt of single continuous sample was extracted from the beginning of each speaker's answers using waveform editing. The length of the samples ranged from 10 -16 s. All utterances were cut at the end of a syntactic phrase or clause that sounds like the end of a unit, normally with falling intonation. Each selected utterance was also low-pass filtered to remove segmental information but to preserve prosodic information such as

stress, rhythm, and intonation. The purpose of including filtered speech was to reduce the impact of segmental and grammatical errors on the accent rating (Munro, 1995). The total stimuli used for accent rating were 100 sentences (2 sentences × 25 speakers × 2 presentations), 25 utterances, and 25 filtered utterances. All stimuli were recorded onto a CD for presentation.

Procedure. The two sentences were presented in two separate blocks. Within each block, the 25 sentences were randomized and presented twice. The 25 utterances and 25 filtered utterances were randomized and presented once in two separate blocks. The listeners rated the sentences first, followed by the utterances and filtered utterances. They performed the rating tasks in small groups of 3-5 in a sound treated language lab where the CD was played through a built-in speaker system at comfortable level. Each stimulus was presented once and the listeners rated each item they heard on an answer sheet by circling a number on a scale of 1 (native-like) to 9 (heavy foreign accent). Before the real rating tasks, the listeners practiced by rating sample sentences, utterances, as well as filtered utterances by both native and nonnative speakers not included in this study. They were encouraged to use the full scale while rating.

Results. Both interrater reliability, (α = .735, .961, .975) and intraclass correlations, (R = .704, .946, .964) for the 11 listeners' ratings ranged from moderate to very high for filtered utterances, sentences, and utterances respectively. There was little difference between the ratings of the two sentences. Therefore, a single mean rating for each speaker's sentences was obtained by averaging the 11 listeners' mean ratings of the 2 sentences. Each speaker's utterance and filtered utterance score was also obtained by taking the listeners' mean ratings.

Figure 1 summarizes the mean ratings by stimulus type for the three speaking groups. The average ratings for the sentence, utterance, and filtered utterance were 5.8, 6.1 and 4.6 for the AP group, 6, 6, and 5 for the CP group, and 1.3, 1.1, and 3 for the EP group respectively. These scores were submitted to a two-way ANOVA with Group (AP, CP, EP) as between subject factor and Stimuli (sentence, utterance, filtered utterance) as repeated measures. The analysis yielded a significant mean effect of Group, $F(2,22) = 84.804$, $p = .000$, a non-significant effect of Stimuli, $F(2,44) = .440$, $p = .647$, and a significant Group × Stimuli interaction $F(4,44) = 9.884$, $p = .000$. A series of one-way between group ANOVAs established

significant differences between the groups on sentence, $F(2,24) = 55.076$, $p = .000$, on utterance, $F(2,24) = 74.948$, $p = .000$, and on filtered utterance, $F(2,24) = 8.667$, $p = .002$. Post hoc (LSD) multiple comparisons revealed no significant differences between the AP and CP groups on any stimulus type. The differences ($\alpha < .01$) were significant between the EP and AP groups and between the EP and CP groups across three stimuli.

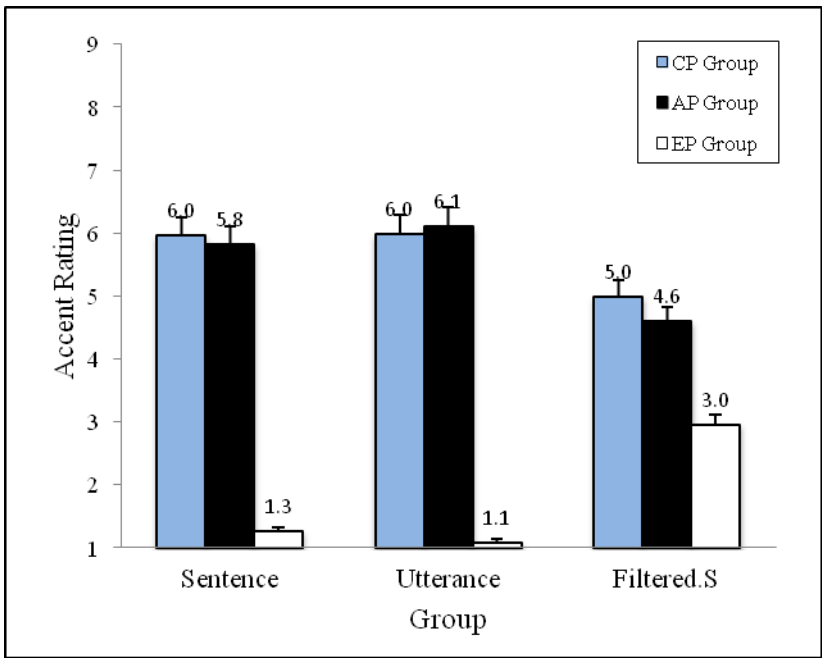

Figure 1. The AP, CP, & EP groups' mean foreign accent ratings given by native English listeners.

A series of paired t-tests for each group on all the possible combinations of stimulus type revealed the Group × Stimuli interaction was due to the differences between the groups on the ratings of the filtered utterance ($\alpha < .05$). For the AP group, the filtered utterance was rated significantly less accented (lower ratings) than the sentence and the utterance. The same held true for the CP Group, except that the difference was significant between the filtered utterance and the utterance but marginal between the filtered utterance and sentence ($p = .059$). The reverse was true for the EP group for which the mean rating of the filtered speech was significantly higher (less native-like) than both the sentence and utterance.

EXPERIMENT 2. ACOUSTIC ANALYSES

This experiment examines native Mandarin speakers' L2 English speech prosodic features in read sentences and spontaneous speech and their relative weights on degree of foreign accent as perceived by native English listeners. In particular, temporal features such as speech rate, articulation rate, pauses, and pause duration will be measured. Pause frequency is the measurement of number of pauses in a speech utterance. Both pause duration and pause frequency have been found as factors that predict perceived degree of foreign accent (Trofimovich & Baker, 2006). The dependent variable is the foreign accent rating scores presented in Experiment 1. The independent variables are speech rate, articulation rate, pause frequency, and pause duration.

Methods

Speakers and Stimuli. The same stimuli, two read sentences and excerpts of spontaneous speech produced by the same speakers in Experiment 1 for foreign accent ratings were used for acoustic measurement.

Measurement. The measurement was carried out on Praat using waveform analyses. For the two sentences, "Most people like to listen to music", and "Do you have a gas cooker in your kitchen", speech rate only was measured for each speaker. Other temporal properties such as pause frequency and pause duration could not be obtained because the experimenter found no such pauses in these short and simple read sentences. For the spontaneous speech, the following measurements were obtained:

1. Speech rate (number of syllables per second);
2. Articulation rate (syllable per second – pauses);
3. Silent pauses (silence of 100 milliseconds and longer) (Kang, 2010, Trofimovich & Baker, 2006, 2007);
4. Mean length of silent pauses (Calculated by counting the number of silent pauses of 100 ms or longer); and
5. Filled Pauses (nonlexical fillers such as *um, uh, er, and so on).*

Results

Sentences. The EP, AP, and CP groups' mean speech rate for each sentence are summarized in Table 1. Overall, for the read sentences, both CP and AP groups had the mean speech rate of 5 syllables per second, about 2 syllable slower than the EP groups' mean, which was 7 syllables per second. There was no noticeable difference between the two sentences. One-way between group ANOVAs on mean speech rate established significant differences between the groups. Post hoc Bonferroni tests (α = .01) established significant differences between the AP and EP groups, and between CP and EP groups, but not between AP and CP groups. Native speakers were two syllables faster than both Mandarin speaking groups in read speech.

Table 1. The EP, AP, and CP Groups' Mean Speech Rate (Syllable/Second) and Standard Deviations for the Two Sentences

Groups	Sentence 1	Sentence 2	Mean
EP	6.7 (0.7)	7.2 (0.8)	7.0 (0.7)
AP	4.5 (0.5)	5.4 (0.6)	4.9 (0.5)
CP	4.5 (0.5)	5.2 (1.0)	4.9 (0.7)

Spontaneous Speech. The results of the acoustic measurements of the spontaneous speech, along with the rating scores, are presented in Table 2. One-way ANOVAs revealed significant differences between the groups on Speech rate, Articulation rate but not significant on Pause frequency, Mean pause duration, & Percentage pause duration. For Speech rate and Articulation rate differences, post hoc Bonferroni tests (α = .01) established significant differences between the AP and EP groups, and between CP and EP groups, but not between AP and CP groups. Native speakers were about 1.5 syllables faster than the Mandarin speakers in spontaneous speech.

Table 2. Means and Standard Deviations of Accent Rating, Speech Rate, Articulation Rate, and Pause Features of Spontaneous Speech by the AP, CP, and EP Groups

Groups	Rating score	Speech rate	Articulation rate	% Pause dur (ms)	# of pauses
AP	6.1 (.72)	3.1 (.91)	3.7 (.90)	1825 (322)	5.7 (3.8)
CP	6.0 (1.0)	2.8 (.54)	3.8 (.47)	2550 (330)	7.4 (2.6)
EP	1.1 (.12)	4.5 (.42)	5.3 (.20)	1600 (202)	4.6 (1.5)

A stepwise regression analysis (adjusted R square = 0.937) with accent rating scores as its dependent variable revealed that when Mean pause duration, Number of pauses and Percent pause duration are controlled, Articulation rate and Speech rate do not contribute to the explanation of accent ratings (see Table 3, Coefficient Table). In other words, when the factors of Speech rate and Articulation rate are removed, the variables of Pauses, Pause duration, and Percentage pause durations are the main predictors of perceived degree of foreign accent. Pause features including number of pauses, mean pause duration, mean % pause duration are predictors of foreign accent in spontaneous speech when speech rate (syl/s, syl/s-pauses) were phrased out.

Table 3. Coefficient Table of the Regression Test

		Unstandardized Coefficients[a]		Standardized Coefficients[b]		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	-1.119	.585		-1.913	.071
	MPauseDur	.007	.001	.452	5.154	.000
	Pauses	.351	.085	.504	4.133	.001
	percentPdur	-.002	.000	-.705	-5.052	.000
	g1	4.830	.304	1.115	15.879	.000
	g2	5.319	.325	1.228	16.384	.000

Note. [a]Dependent Variable: rating. [b]Dependent Variable: rating.

DISCUSSION AND CONCLUSIONS

The goal of the current study was to investigate which speech temporal properties predict native Mandarin speakers' degree of foreign accent as perceived by native English listeners and whether advanced native Mandarin speakers with over 12 years of LOR in North America had an advantage over those with zero LOR in producing more native like speech temporal features that may lead to reduced degree of foreign accent. Revisiting the research questions 1 that asked in what ways native Mandarin speakers produced different English prosodic temporal features as compared with native English speakers, the results showed that Mandarin speakers' speech rate was significantly slower than the native English speakers in both read sentences and spontaneous speech. This difference was consistent across both native Mandarin speaking groups:

both AP and CP groups read the sentences at the rate of 5 syllable per second, 2 syllables slower than the EP group, which had the mean rate of 7 syllable per second.

For spontaneous speech, The AP and CP groups' production rate was about 3 syllables per second. Compared with the native speakers' speech rate of 4.5 syllable per second, the Mandarin speakers were about 1.5 syllable slower in spontaneous speech. This rate was consistent with and without the pauses (Articulation rate). There was no significant difference between the AP and CP groups in speech rate despite the differences in 12 years of LOR. It is obvious 12 years of LOR for the AP group did not seem to give them any significant advantage over the CP group in terms of producing native like speech temporal features in their L2 production. As shown in Experiment 1, the AP and CP groups received the almost the same rating scores by native English listeners for both read sentences and spontaneous speech. The measurements of speech temporal features, reflect, or, are in agreement with, the foreign accent rating scores both AP and CP groups received in Experiment 1.

The second research question asked whether native Mandarin speakers' speech temporal features contribute to perceived degree of foreign accent and if so, which temporal features are the predictors. The results of the spontaneous speech showed that pause features including number of pauses, pause duration, and percentage pause duration are the predictors of foreign accent. It is interesting to note that though speech rate and articulation rate differences between the native and nonnative groups are significantly different for the spontaneous speech (Native English speakers were 1.5 syllable faster than the Mandarin speakers), they are not the predictors of degree of foreign accent. Once these two variables are phased out, the pause features all became predictors. The results suggest that raters are mostly influenced by how often or how badly a speaker pauses, rather than how fast he/she speaks. One explanation could be that natural native-like speeches, which do not have awkward or untimely pauses, are given higher ratings than speeches that have frequent pauses or pauses at the wrong places. From a related perspective, speeches that do not have unnecessary pauses are faster. Once pauses are taken into account, speed of the speech no longer matters to the raters.

The findings of pause features as predictors of Mandarin speakers' degree of foreign accent are partly in agreement with Trofimovich and Baker's (2006) study which concluded that both speech rate and pause duration are predictors of native Korean speakers' L2 English foreign accent. Yet, not in agreement with their study is that speech rate is not a

predictor of Mandarin speakers' foreign accent in the current study. Trofimovich and Baker's (2006) study on native Korean speakers also concluded that age of arrival (AOA) influenced the speech rate, pause frequency, and pause duration. One limitation of the current study is that it examined the LOR effect but not the AOA factor on L2 speech prosody and accent rating.

Another limitation of this study is that the acoustic measurements were limited to speech temporal features. It is not known how other prosodic features such as stress timing and pitch range or movement contribute to perceived degree of accent by these two speaking groups. Future studies need to include these suprasegmental factors as well as segmental features as variables for a more complete picture of the predictors of foreign accent.

Overall, Mandarin speakers, both AP and CP groups, produced English L2 speech at a significantly slower rate than the native English speaking group. Yet, speech rate and articulation rate of spontaneous speech are not the predictors of accent rating scores. In contrast, pause features, including pause frequency and pause duration are the predictors of perceived degree of foreign accent for native Mandarin speakers. Over 12 years of LOR in North America did not appear to give the AP group any significant advantages in producing native like speech temporal features and in reduced degree of foreign accent.

REFERENCES

Chen, Y., Robb, M. P., Gilbert, H. R., & Lerman, J. W. (2001). A study of sentence stress production in Mandarin speakers of American English. *The Journal of the Acoustical Society of America, 109*(4), 1681-1690.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84*, 70-79.

Flege, J. E., Munro, M. J., & MacKay, I. (1995). Factors affecting strength of perceived foreigh accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. [doi: DOI: 10.1016/j.system.2010.01.005]. *System, 38*(2), 301-315.

Kang, O., Rubin, D. O. N., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal, 94*(4), 554-566.

MacKay, I. R. A., Flege, J. E., & Imai, S. (2006). Evaluating the effects of chronological age and sentence duration on degree of perceived foreign accent. *Applied Psycholinguistics, 27*, 157-183.

Munro, M. J. 1995. Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition, 17, 17-34*.

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech, 38*, 289-306.

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

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191-215.

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

Riney, T. J., Takagi, N., & Inutsuka, K. (2005). Phonetic parameters and perceptual judgments of accent in English by American and Japanese listeners. *TESOL Quarterly, 39*, 441-466.

Trofimovich, P., & Baker, W. (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, & Baker, W. (2007). Learning prosody and fluency characteristics of second language speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics, 28*(2), 251-276.

Wang, X., & Chen, J. (2007). Effects of length of residence and speech activities on degree of foreign accent. In J. Trouvain & W. Barry (Eds.) *The Proceedings of the 16th International Congress of Phonetic Sciences*, (pp. 1725-1728). Saarbrucken: Germany.

Zhang, Y., Nissen, S. L., & Francis, A. L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America, 123*(6), 4498-4513.