A MANDARIN SPEAKER'S INTONATIONAL EMPHASIS IN ENGLISH AND MANDARIN LECTURES

Di Liu, Boston University

Research suggests that native English speakers and non-native English speakers use different English intonation patterns (Pickering, 2004; Wennerstrom, 1998) and that prosody significantly affects English learners' intelligibility and comprehensibility (Derwing, Munro, & Wiebe, 1998). However, cross-linguistic studies of non-native English speakers' prosody production in English and in their first languages are highly limited. This study investigates the given-new stress connection (GNSC) realized in a Mandarin speaker's speech in English and in the same speaker's speech in Mandarin. The results show that the Mandarin speaker lacks the ability to effectively use pitch as a prosodic cue to mark new versus old information in English but the same speaker is able to effectively use prosodic cues to contrast new versus old information in Mandarin. Furthermore, a large portion of the new information that should be emphasized in English is emphasized in the Mandarin version but not the English version of the lecture. The findings of this study imply that Mandarin speakers may use prosodic features to emphasize new information while speaking in Mandarin but they may not fully transfer this strategy to English. The findings of this study also imply that ESL learners' first language could be a valuable resource in English prosody teaching.

INTRODUCTION

Pronunciation teaching has changed drastically in recent years (Murphy & Baker, 2015). One profound change in the contemporary field of pronunciation pedagogy is a shift in the features upon which teachers focus on in their classrooms. Pronunciation was dominated by the teaching of segmental features for many years. In the past 30 years, however, a growing number of researchers and teachers realized the significant influence that prosody has on intelligibility and comprehensibility; this led them to call for a shift of focus towards prosodic features (Anderson-Hsieh, Johnson, & Koehler, 1992; Celce-Murcia, Brinton, & Goodwin, 2010; Derwing, Munro, & Wiebe, 1998).

PROSODY

Often described as the "music" or "melody" of language (Allen, 1971; Wennerstrom, 2001), prosody encompasses a range of linguistic variables including intonation, stress, pauses, and rhythm. Researchers generally agree that prosody encodes meanings and has various pragmatic

functions (Levis & Wichmann, 2015; Pierrehumbert & Hirschberg, 1990; Reed & Michaud, 2005; Wennerstrom, 2001). At the lexical level, prosody can change the categories of lexical items, for example, *REcord* with lexical stress on the first syllable is a noun and refers to something permanent and *reCORD* with lexical stress on the second syllable is a verb and means to convert something into permanent form.

At the phrasal level, prosody has much broader functions and a less fixed influence on meaning. A sophisticated understanding of prosodic elements is important because they mark information structure, including the distinction between new and old constituents (Hahn, 2004; Pickering, 2004; Pierrehumbert & Hirschberg, 1990; Wennerstrom, 2001). At the phrasal level, prosody shifts the focus of utterances and can mark contrastive stress. Stress can change not only the meaning of a sentence, but also its focus and implication. For example: *It is not MY responsibility to do this job* and *It is not my RESPONSIBILITY to do this job* realize focus on different constituents, and the implications differ due to stress alone. The former sentence suggests that it is someone else's responsibility to do the job, whereas the latter emphasizes the fact that the speaker is not obliged to do the job. Hahn (2004) demonstrated that sentence level primary stress plays an important role in intelligibility by investigating native speakers' reaction times, memory and attitudes to sentences read by an international teaching assistant "with primary stress correctly placed (e.g., *The URban environment is more individualistic than the RUral environment.*), misplaced (e.g., *The urban environment is more individualistic than the rural enVIronment.*) or missing (e.g., *The urban environment is more individualistic than the rural environment*)" (p. 206). She found that when the primary stress was correctly placed, native listeners processed the lecture more easily, recalled significantly more information, and provided significantly better speaker evaluations.

PROSODY FUNCTIONS IN CLASSROOM

Researchers have noted the pragmatic functions of prosody in classroom. Chun (1988) states that, in classrooms, teachers use a wide range of communicative functions including, "addressing others, selecting the next speaker, choosing the topic, interrupting, asking for clarification, changing the subject, and concluding a discussion" (p. 82). Hellermann (2003) reviewed over 25 hours of classroom IRF (initiation-response-feedback) interaction and confirmed the communicative value of prosody in a classroom. His analysis shows that teachers and students systematically use intonation in creating an effective classroom discourse and that teachers use complex prosody packaging while providing feedback to students. Wennerstrom (1998) proposed that there is an intonation system in English that functions at the discourse level to signal relationships in information structure. She proposes a model in which intonation functions as a grammar of cohesion.

ITAS' CLASSROOM PROSODY

Pickering compared native speaker teaching assistants' (NSTAs') and international teaching assistants' (ITAs') tone choice and use of intonation in the classroom. She argues that, "intonation bears a high communicative load in terms of information structuring and rapport building between discourse participants" (2001, p. 234). Pickering's (2004) subsequent analysis of NSTAs' instructional monologues reveals a hierarchy of phonologically defined units that coincide with structural boundaries at other levels of discourse; these prosodic elements contribute to the overall organization of the teaching presentations, whether the instructors recognize it or not. Her analysis of parallel ITA data, however, demonstrates that Mandarin-speaking ITAs are generally incapable of effectively controlling this level of structural organization. Specifically, ITAs in her studies made use of an overall narrower pitch range and were unable to consistently manipulate *key*, which is defined by Brazil (1981) as the starting pitch level of an intonation unit, and *tone choices* to create intonational paragraphs.

PROSODY TEACHING

Prosody is hard to teach due to its fluid nature and dependence on contextual information. Previous literature that intends to generalize the "meanings" of pitch contours concludes that the "meanings" are highly context-dependent and that even native speakers can hardly reach a consensus of the "meanings" of pitch contours (Lieberman, 1967, p. 124). Native English speakers gain this contextual information and knowledge over many years and through frequent communication with other native speakers. However, it is challenging for non-native speakers of English to gather this contextual information in a much shorter learning period with much less communication with native speakers of English.

One possible solution to the issue regarding contextual information is to transfer certain prosodic functions from one language to another. If we consider the situation for Mandarin learners of English, there is a relative lack of empirical studies that directly compare Mandarin speakers' English and Mandarin use of similar functions across their two languages. It is not until recently that researchers have started to investigate cross-linguistic functions of prosody between Mandarin and English.

MANDARIN PROSODY

Ouyang and Kaiser (2015) conducted a study to investigate corrective focus and given/new contrast in Mandarin prosody. The results show that even a language with lexical tones (Mandarin), uses prosodic features (i.e., fundamental frequency (F0), duration and intensity) that are used to encode information structure in English. In fact, not only can prosodic cues indicate importance at the discourse level, they also distinguish different types of information structure in Mandarin. Analyzing telephone dialogue data, Ward, Li, Zhao and Kawahara (2016) conclude "a significant fraction of the prosody of Mandarin dialog relates to pragmatic functions", which

include, but are not limited to "back-channeling, turn yielding, turn starts, negative evaluation, bids for empathy, topic progression, topic exhaustion, and contrast" (p. 1234). Of course, it is still possible that even though Mandarin and English both use the same suprasegmental features in realizing the same functions, they do so to different degrees. However, a recent study presented by Ip and Cutler (2016) suggests that pitch levels used to mark focus in Mandarin are even higher than they are English when marking focus. Comparing five different types of prosodic focus by native English speakers and Mandarin speakers, Ip and Cutler find that "native speakers of Mandarin resemble English speakers in their tendency to signal focus by manipulation of duration, pitch range, and intensity." (p. 333) Furthermore, they find that "Mandarin speakers reliably show greater increase in pitch for new-information focus (either as pitch range or mean/maximum)" (p. 333). Although these studies suggest that there are similarities between Mandarin and English prosody, studies that directly compare the same speaker's prosody production in English and Mandarin are lacking.

GNSC EFFECT

Among all the prosodic features, stress may be the most important and widely-studied one. Stress constituents usually have a higher pitch level, longer duration, greater amplitude and full vowels. One particular function of stress that warrants attention is the "GNSC" effect. Pierrehumbert and Hirschberg (1990) argue that in English, new and contrastive information tends to be realized with a high pitch accent, and old or given information tends to be realized with a low pitch accent. Hahn (2004) refers to this relationship as the given-new stress connection (GNSC). The present study investigates if the GNSC is demonstrated in a Mandarin speaker's speech in English and if the same relationship is shown in the speaker's speech in Mandarin. The research questions are:

1.   How does a speaker who speaks both Mandarin and English demonstrate the given-new stress connection (GNSC) in Mandarin and English?

2.   To what degree is the given-new stress connection (GNSC) realized in a native Mandarin speaker's Mandarin lecture?

3.   To what degree is the given-new stress connection (GNSC) realized in a native Mandarin speaker's English lecture?

METHOD

The material analyzed is a massive open online course (MOOC): *Principle of Electric Circuits* retrieved from EdX. This course is delivered in two languages (Mandarin and English) by the same speaker, who is a native speaker of Mandarin and who was a visiting scholar at the Massachusetts Institute of Technology. The English version is a direct translation of the Mandarin version, and the speaker made little modification of the content while he was giving

the lecture using PowerPoint slides. The recordings were made for the same purpose and the same online platform, so that the qualities of the sound files are similar.

The lectures were downloaded from EdX and analyzed and annotated using the speech analysis software *Praat*. The speech samples are annotated based on the ToBI (Tone and Break Indices) conventions (Beckman & Hirschberg, 1994). The English version of the lecture was annotated based on the MAE (Mainstream American English)_ToBI conventions (see Figure 1).
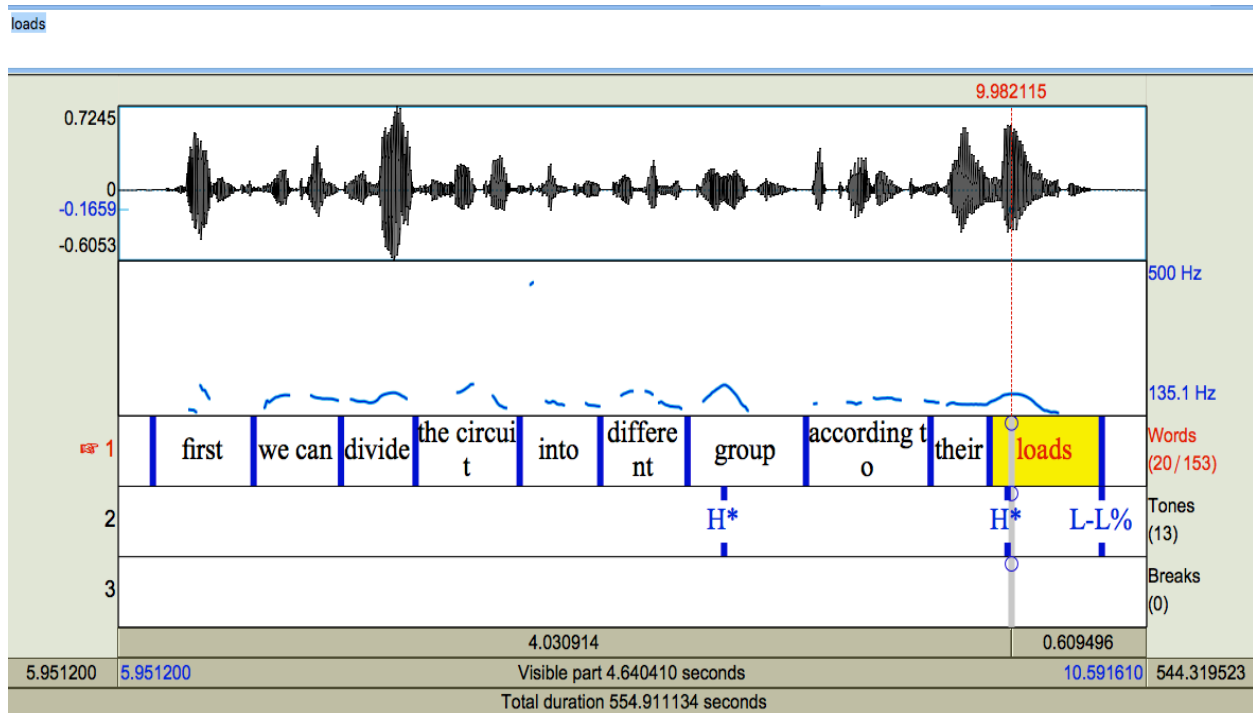

Figure 1. Sample MAE_ToBI Transcription.

The Mandarin version of the lecture was annotated according to the Pan Mandarin_ToBI conventions (see Figure 2).
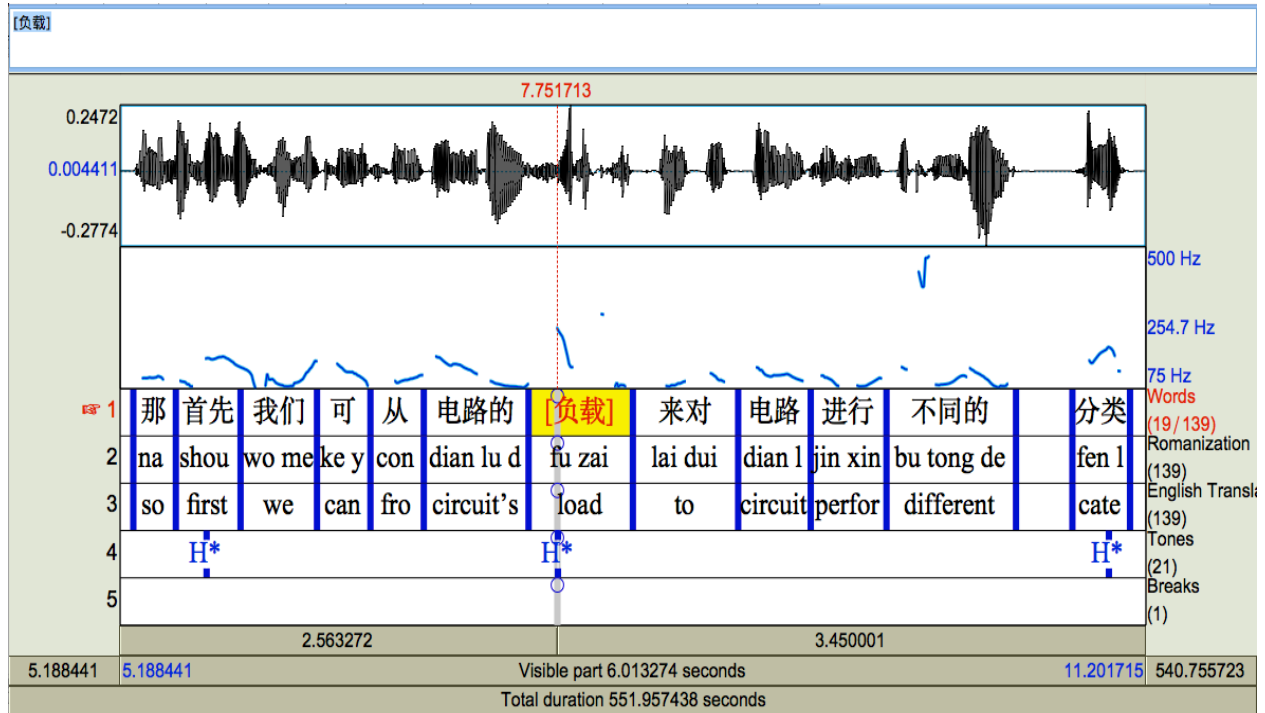
Figure 2. Sample Pan Mandarin_ToBI Transcription.

Because the goal of this research is to analyze items that were realized as both "new" and given," 32 concepts that were mentioned more than one time are analyzed below. If the concept is mentioned the first time in the lecture, it was coded as "new" and if the concept has been mentioned before in the lecture, it was coded as "old" (e.g. load $_{new}$ vs. load $_{old}$). The maximum pitch level of these words and phrases were retrieved using the built-in pitch elicitation function in *Praat*. Paired t-tests were conducted to compare the maximum pitch level across different groups.

RESULTS

A paired t-test was conducted to determine if there was a difference in the mean maximum pitch level of the Mandarin words/phrases as compared to the mean maximum pitch level of the English words/phrases. The total number of words/phrases analyzed is 64 for each group (32 new concepts and 32 old concepts). The Mandarin version of the lecture has a higher mean pitch level (M = 176.25) as well as a higher standard deviation (SD = 53.49) compared to the mean pitch level (M = 160.99) and the standard deviation (SD = 30.42) of the English version of the lecture, suggesting that the Mandarin speaker uses a higher pitch level and that his pitch fluctuates more when he speaks in Mandarin than when he speaks in English. A paired t-test was conducted to determine if this difference was statistically significant (t = 2.0362, df = 63, p = 0.04594). The null hypothesis is rejected, and we can conclude that there is a statistically significant difference between the pitch level of words/phrases in the speaker's Mandarin speech

and English speech and the speaker's Mandarin speech has a significantly higher pitch level compared to the same speaker's English speech.
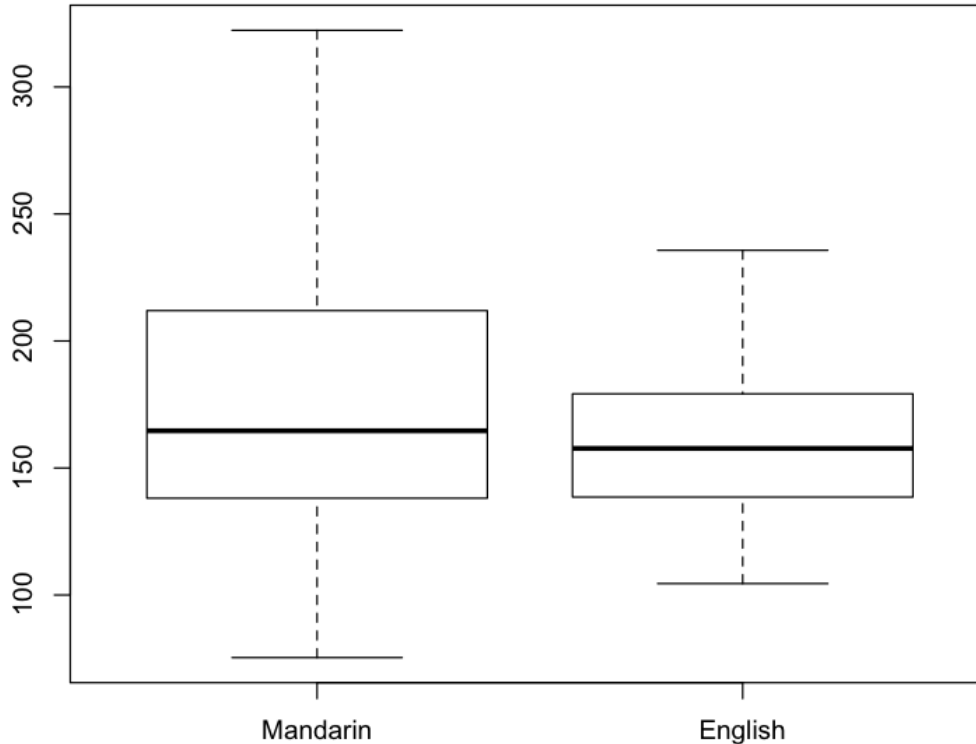

Figure 3. The speaker's mean maximum pitch levels (in Hz) in Mandarin and English.

A paired t-test was conducted to determine if there was a difference in the mean maximum pitch level of the new concepts as compared to the mean maximum pitch level of the old concepts in the Mandarin lecture. The total number of words/phrases analyzed is 32 for each type of information. The mean and standard deviation of new information (M = 195.09, SD = 47.32) are significantly different from those of the old information (M = 157.42, SD = 47.32). A paired t-test determined that this difference was statistically significant (t = 3.0707, df = 31, p = 0.0044). The null hypothesis is rejected, and we can conclude that there is a statistically significant difference between the pitch level of the new information and old information in the speaker's Mandarin speech. The pitch level of the new information in the speaker's Mandarin speech is significantly higher than the pitch level of the old information in the same speaker's Mandarin speech.

Figure 4. The mean maximum pitch levels (in Hz) of Mandarin old and new information

A paired t-test was also conducted to determine if there was a difference in the mean maximum pitch level of the new concepts as compared to the mean maximum pitch level of the old concepts in the English lecture. The total number of words/phrases analyzed is 32 for each type. The mean and standard deviation of new information (M = 164.83, SD = 28.16) are close to those of the old information (M = 157.16, SD = 32.52). A paired t-test was conducted to determine if this difference was statistically significant, and it was not (t = 1.4744, df = 31, p = 0.1505). This lack of statistically significant difference between the pitch levels of new and old information in the English lecture suggests that GNSC effect in the Mandarin speaker's English speech is not significant.
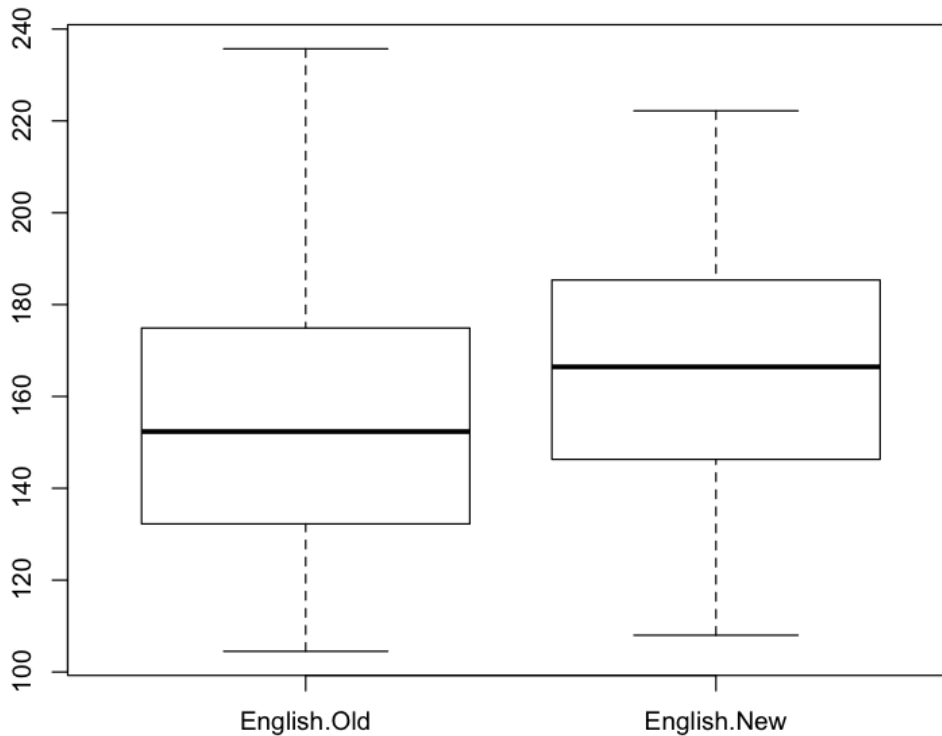
Figure 5. The mean maximum pitch levels (in Hz) of English old and new information

CONCLUSION

This study investigates the given-new stress connection (GNSC) realized in two versions (English and Mandarin) of the same lecture given by a native Mandarin speaker who is fluent in both English and Mandarin. The results show that the Mandarin speaker uses a higher pitch level and a more varied pitch contour while delivering the lecture in Mandarin. In terms of the GNSC, the Mandarin speaker demonstrates significantly different pitch levels between new and old information in the Mandarin version of the lecture but does not demonstrate statistically significantly different pitch levels when he was delivering the same lecture in English. This result suggests that the Mandarin speaker lacks the ability to effectively use prosodic cues to contrast new versus old information in his English lecture even though the speaker is able to use prosodic cues to contrast new versus old information effectively when giving the same content in Mandarin. In other words, the Mandarin speaker does not fully transfer his skills and strategies of using prosodic cues from Mandarin to English.

Students' lack of awareness of the importance of prosody hinders prosody learning. As Gilbert (2014) said, "[w]hile it might seem self-evident to a teacher that intonation is important, students

may simply consider imitating pitch contours 'awkward,' even if they seldom tell teachers that they 'feel silly' speaking in this way. As a result of these differing perspectives, students "may walk out of the class without having accepted the system at all" (p.125). Gilbert's observation echoes what Allen (1971) noted 40 years ago: "there is little carry-over into the students' own conversations outside the classroom and the listen and repeat approach has never yielded satisfactory long-term results" (p. 79). The fact that this issue remains unsolved after so many years even though researchers have acknowledged it suggests that we may lack effective techniques to enhance learners' awareness of the importance of English prosody. The findings of this study, however, suggest a possible route. If students understand that they have been using prosody to realize a range of pragmatic functions in their first language(s) and that English uses prosody for the same purposes, it may be much easier for them to accept the system and adopt English prosody.

This study also finds that even two languages (Mandarin and English) that differ significantly in lexical level prosody (i.e., Mandarin is a tone language and English is a non-tone language) share similar prosodic functions. Therefore, based on the findings of this study, English learners' L1s might be a valuable resource in prosody teaching. Pronunciation teachers might be able to improve the intelligibility and comprehensibility of the learners by facilitating positive prosodic transfer between students' L1s and English discourse prosody functions. Future studies should investigate prosodic functions in other languages and compare the relatively well-studied English prosody functions with the less-studied prosody functions of other languages. Finally, the findings of this study are based on the analysis of a single subject's speech therefore might lack generalizability. Future studies should include more participants to confirm or contradict the findings of this study.

ABOUT THE AUTHOR

Di Liu is a third-year doctoral student in Boston University. He is working with Dr. Marnie Reed on prosody research and teaching. His research includes (1) cross-linguistic comparison and transfer of Mandarin and English classroom prosody; (2) asymmetrical cognitive load imposed by processing native and nonnative speech and (3) constraints on downtrend as a function of tone inventory structure: evidence from Cantonese and Mandarin. Serving as the TA of a Computer-Assisted Language Learning course in BU for the past three years, he is also designing a curriculum/textbook: Technology-Enhanced Pronunciation Teaching, which intends to provide a systematic pedagogical application of various software/websites/apps that are available to teachers.

520 Main Street, Apt. 901
Malden, MA, 02148
917-868-5796
diliu@bu.edu

REFERENCES

Allen, V. (1971). Teaching intonation, from theory to practice. *TESOL Quarterly, 5*(1), 73-81.

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

Beckman, M. E, & Hirschberg, J. (1994). The ToBI annotation conventions. Online MS. Available at http://www.ling.ohio-state.edu/tobi/ame.tobi/annotation.conventions.html

Brazil, D. (1981). The place of intonation in a discourse model. In R. M. Coulthard & M., Montgomery (Eds.), *Studies in Discourse Analysis*. (pp. 146-157). London: Routledge & Kegan Paul.

Celce-Murcia, M., Brinton, D., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide*. New York: Cambridge University Press.

Chun, D. (1988). Teaching intonation as part of communicative competence: Suggestions for the classroom. *Die Unterrichtspraxis / Teaching German, 21*(1), 81-88.

Derwing, T. M., Munro, M. J & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48,* 393-410.

Gilbert, J. (2014). Myth 4: Intonation is hard to teach. In L. Grant, (Ed.), *Pronunciation myths: Applying second language research to classroom teaching*. Ann Arbor, MI: University of Michigan Press.

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2), 201-223.

Ip, M. H. K., & Cutler, A. (2016). Cross-language data on five types of prosodic focus. *Proceedings of Speech Prosody 2016*, May 31-Jun 3, Boston

Hellermann, J. (2003). The interactive work of prosody in the IRF exchange: Teacher repetition in feedback moves. *Language in Society, 32*(1), 79-104.

Levis, J. M. & Wichmann, A. (2015). English intonation: Form and meaning. In M. Reed & J. Levis (Eds.). *The Handbook of English Pronunciation* (pp. 139-155), West Sussex, UK: John Wiley & Sons, Inc.

Lieberman, P. (1967). *Intonation, perception, and language* (M.I.T. research monograph 38). Cambridge: M.I.T. Press.

Murphy, J. & Baker, A. (2015). History of ESL pronunciation teaching. In M. Reed & J. Levis (Eds.). *The Handbook of English Pronunciation* (pp. 36-65), West Sussex, UK: John Wiley & Sons, Inc.

Ouyang, I. C. & Kaiser, E. (2015). Prosody and information structure in a tone language: An

investigation of Mandarin Chinese. *Language, Cognition and Neuroscience, 30*(1-2), 57-72.

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

Pickering, L. (2004). The structure and function of intonational paragraphs in native and nonnative speaker instructional discourse. *English for Specific Purposes, 23*(1), 19-43.

Pierrehumbert, J. & Hirschberg J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R., Cohen, J. Morgan, & M. E. Pollack (Eds.) *Intentions in Communication* (pp. 271-323). Cambridge, Massachusetts: The MIT Press.

Reed, M., & Michaud, C. (2005). *Sound Concepts: An Integrated Pronunciation Course*. New York: McGraw-Hill.

Skidmore, D., & Murakami, K. (2010). How prosody marks shifts in footing in classroom discourse. *International Journal of Educational Research, 49*(2-3), 69-77.

Ward, N. G., Li, Y., Zhao, T., & Kawahara, T. (2016). Interactional and pragmatics-related prosodic patterns in Mandarin dialog. *Proceedings of Speech Prosody 2016*, May 31-Jun 3, Boston.

Wennerstrom, A. (1998). Intonation as cohesion in academic discourse: A study of Chinese speakers of English. *Studies in Second Language Acquisition, 20*(1), 1-25.

Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*. Oxford; New York: Oxford University Press.