

PRESENTATION/POSTER

ASYMMETRICAL COGNITIVE LOAD IMPOSED BY PROCESSING NATIVE AND NON-NATIVE SPEECH

Di Liu, Boston University

Marnie Reed, Boston University

Intonation affects information processing and comprehension. Previous research has found that some international teaching assistants (ITAs) fail to exploit English intonation, potentially posing processing difficulties to students who are native English speakers. However, researchers have also found that non-native listeners found it easier to process sentences given by a non-native speaker with a shared language background, leading to an interlanguage speech intelligibility benefit (ISIB). Therefore, how native speaker teaching assistant (NSTA)'s and ITA's classroom speech affects the processing, comprehension, and attitudes of listeners with different language backgrounds needs to be further investigated. Using a dual-task paradigm, a comprehension questionnaire, and an attitudinal questionnaire, the present study investigates how the pronunciation and intonation of a NSTA and an ITA affect native English speakers' and Mandarin-speaking English learners' processing and comprehension of a lecture, and attitudes towards the two instructors. The present study found shared processing advantages when the listeners shared the L1 of the speaker, but overall lecture comprehension and attitude were unaffected. These findings support and extend prior research studies surveying ITAs' intonational patterns and ISIB. These findings also have implications for research on the teaching of English pronunciation to non-native instructors.

INTRODUCTION

The field of English pronunciation teaching has witnessed two profound and significant changes in the past few decades. First, scholars and teachers rejected accent-reduction oriented pronunciation teaching and embraced the intelligibility principle. Second, more attention has been given to suprasegmental features such as intonation, stress, pauses, and rhythm.

Intelligibility, comprehensibility, and accentedness

Munro and Derwing (1995) conducted a study in which they asked native English speakers to listen to English sentences produced by native Mandarin speakers, transcribe the sentences, and rate each speaker's comprehensibility and accentedness. The result shows that, although it took significantly longer for the listeners to verify whether a message is correct, errors in the verification and transcription of the Mandarin-accented utterances were fairly rare (5-10%). Based on this and prior work by Abercrombie (1949), Munro and Derwing introduced three distinctly measurable concepts to second language pronunciation: Intelligibility, Comprehensibility and Accentedness. Intelligibility enhancement then became the new goal of pronunciation teaching. The following table from Derwing and Munro's (2005) study summarizes their definition and measurement of these three concepts.

Table 1

Intelligibility, comprehensibility and accentedness (Derwing & Munro, 2005, p. 385)

Intelligibility, Comprehensibility, and Accentedness		
Term	Definition	Measure
Intelligibility	The extent to which a listener actually understands an utterance	Transcription task % words correct
Comprehensibility	A listener's perception of how difficult it is to understand an utterance	Scalar judgment task 1 = extremely easy to understand 9 = extremely difficult to understand
Accentedness	A listener's perception of how different a speaker's accent is from that of the L1 community	Scalar judgment task 1 = no accent 9 = extremely strong accent

While the transcription task may provide a relatively accurate measurement of intelligibility, Munro and Derwing acknowledged the complexity of intelligibility in their later studies. In particular, Munro and Derwing (2015) argued that “although the focus in L2 research is often on the characteristics of L2 speakers’ productions, listeners play a crucial role in establishing the consequences of those characteristics” (p. 388). The current study, therefore, investigates how a NSTA’s and a Mandarin-speaking ITA’s classroom speech affects listeners who are native English speakers and Mandarin-speaking English learners. This design partially simulates university classrooms in which the instructors and the students may or may not share the same language background. The results of the present study, therefore, can help us to understand how instructors’ pronunciation affects the processing, comprehension, and attitudes of students with different language backgrounds in a classroom setting.

Munro and Derwing also distinguish between local intelligibility— which refers to “how well listeners recognize relatively small units of speech, such as segments and words, outside of a larger meaningful context” (p. 381)— and global intelligibility— which entails “larger units of language that include rich contextual information” (p. 381). They argue that instructors whose goal is to enhance students’ communicative competence should focus on global intelligibility, though local intelligibility proves more important for understanding the L2 acquisition process. Using a fairly long (742-word) passage delivered in the format of an online course, the present study intends to provide implications for NSTAs’ and ITAs’ global intelligibility within a classroom discourse context.

Suprasegmental features, prosody, and intonation

The terms “suprasegmentals” and “prosody” are commonly used interchangeably. They encompass a number of linguistic features such as intonation, pauses, stress, and rhythm. Pronunciation was formerly dominated by segmental teaching focusing on correct production of

individual consonant and vowel sounds. However, research in the past thirty years clearly demonstrated that, in English, the suprasegmental features, or prosodic features, also play a crucial role in intelligibility (Anderson-Hsieh, Johnson, & Koehler, 1992, Derwing, Munro, & Wiebe, 1998). Gilbert (2014), for example, argued that “[i]n English, prosodic cues serve as navigation guides to help the listener follow the intentions of the speaker. These signals communicate emphasis and make clear the relationship between ideas (new and old information) so that listeners can readily identify these relationships and understand the speaker’s meaning” (p. 123).

Within the last two decades, a growing number of researchers working on prosody have narrowed their scope and shifted their attention to the function of prosody in academic settings. Researchers have argued that prosody serves as a contextual cue and organizational device in classroom interaction and that teachers may choose different configurations of prosodic resources to help embody the type of interaction they want to accomplish and to control the direction of classroom discourse (Hahn, 2004; Hellermann, 2003; Pickering, 2004; Wennerstrom, 2001). For example, Chun (1988) states that, in classrooms, “teachers use a much wider range of communicative functions; their ‘privileges’ include addressing others, selecting the next speaker, choosing the topic, interrupting, asking for clarification, changing the subject, and concluding a discussion” (p.82). Wennerstrom (1998) proposed that the intonation system in English functions at the discourse level to signal relationships in information structure and mark interdependencies among constituents; she proposes a model in which intonation functions as a grammar of cohesion.

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. Furthermore, he found that teachers use complex prosody packaging while providing feedback to students. Skidmore and Murakami (2010) investigated an additional, important function of prosody in classroom discourse; specifically, in teacher-led dialogue, prosody might serve to signal shifts from one kind of teacher-student interaction to another. They found that prosody signals the boundaries between multiple teacher-led IRF activities, such as the passage from “thinking aloud” to an authoritative discourse used when assigning homework.

ITAs’ pronunciation

Having compared NSTAs’ and ITAs’ intonation in classroom speech, Pickering (2001) argued that “intonation bears a high communicative load in terms of information structuring and rapport building between discourse participants” (p. 234). Pickering’s (2004) 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 lack effective control of English prosody at this level of structural organization. Specifically, the Mandarin speaking International teaching assistants (ITAs) “tend to create a flat, monotonic pitch structure unfamiliar to NS hearers” (Pickering, 2001, p. 249). ITAs’ failure in exploiting English prosody may affect English listeners’ processing of classroom speech. Gilbert (2014) noted that “emphasis that conveys the wrong meaning, or thought groups that either run together or break in inappropriate

places, cause extra work for the listener who is trying to follow the speaker's meaning. If the burden becomes too great, the listener simply stops listening" (p.123).

Using a dual-task paradigm in which subjects were asked to complete a visual task while listening to a lecture, Hahn (2004) found that the ITA's correct use of sentence stress helped native speakers of English to process the information more easily. Native English speakers also recalled significantly more content and rated the speaker more favorably when the sentence stress was correctly placed. Hahn's groundbreaking study established a foundation for extending the investigation to the discourse level, and to the effects of ITAs' speech on non-native English-speaking listeners.

Interlanguage speech intelligibility benefit (ISIB)

The characteristics of ITAs' classroom speech, although negatively affecting English listeners' processing, may not pose the same challenge for listeners who are non-native speakers of English. For example, Bent and Bradlow (2003) found that non-native listeners considered speech from a high proficiency non-native speaker as intelligible as speech from a native speaker. They argued that there is an "interlanguage speech intelligibility benefit" (ISIB) between speaker and listener who share the same L1. They also found that ISIB exists even when the non-native listeners and speakers share different L1s. However, their study focuses on sentence level intelligibility; whether ISIB exists at the discourse level and in classroom discourse needs to be further investigated. Using a dual-task paradigm similar to the one used by Hahn (2004), the present study investigates how the pronunciation and intonation of a NSTA and an ITA affect processing, comprehension, and attitude of the listeners who are native English speakers and Mandarin-speaking English learners.

Research questions

The research questions are:

1. How do the pronunciation and intonation of a NSTA and an ITA affect listeners' processing and comprehension of a lecture and their attitudes toward the speakers?
2. How do listeners' L1s affect their processing and comprehension of a lecture given by an NSTA and an ITA and their attitudes toward the instructors?

METHODS

Participants

The participants are twenty-one undergraduate and graduate students recruited in a university in the US. Nine participants are native speakers of English; twelve are Mandarin-speaking English learners.

Materials

One NSTA and one native Mandarin-speaking ITA were recruited to be the speakers of this study. The two teaching assistants are both female doctoral students. The ITA is a high proficiency English speaker with a 2013 TOEFL iBT score over 100 and a U.S. master's degree. When she was asked to be a speaker of the present study, she had been living in the US for six years.

The two teaching assistants were asked to “teach an online mini lesson” with the same lecture script adapted from *Sound Concepts*, a pronunciation textbook written by Reed and Michaud (2005). A paragraph of this lecture is given below:

Okay, today we're gonna talk about the universality of human emotions. First of all, let me say that this theory is attributed to Paul Ekman, a professor of psychology who's known as “the world's most famous face reader.” Dr. Ekman's based at the University of California Medical School at San Francisco, but he's done research all over the world. Dr. Ekman says he's always been interested in emotions, ever since he was a teenager. And, being a photographer since he was twelve, he just naturally decided to look at facial expressions. In Ekman's view, it turns out there're seven basic human emotions: anger, sadness, fear, surprise, disgust, contempt, and happiness. All of these emotions have clear facial signals. There're actually 43 facial muscles that combine to reveal these emotions...

Procedure

Participants were randomized into two groups. One group listened to the lecture given by the NSTA, the other group listened to the lecture given by the ITA. Using a “dual task paradigm” to investigate the processing of the lecture by the listeners, participants were asked to accomplish two tasks simultaneously—listen to a lecture, and indicate the orientation of projected images. Based on cognitive capacity, the faster the participants complete the orientation task, the easier it is to process the lecture content.

Superlab, a presentation software, was used to deliver the lectures. While the participants were listening to the lecture, they saw images that appeared every five seconds on the screen. They were asked to press either of two keys to indicate for each image whether it was upright (press “j”) or inverted (press “f”) as quickly as possible. The participants were instructed that their primary task was to comprehend and remember the content of the lecture, and their secondary task was to press the correct key as quickly as they can.

After listening to the lecture, participants were asked to complete seven open-ended comprehension questions. The questions tested participants' comprehension of the main idea and details in the lecture (e.g., Why did Ekman go to Papua New Guinea?)

At the end of the study visit, participants were asked to complete a speaker evaluation questionnaire adapted from Hahn (2004). The participants were asked to rate the instructors based on a 5-point Likert scale. Two sample questions are provided below:

Q2: How would you characterize the instructor's ability to explain?

Excellent 5 4 3 2 1 Very poor

Q5: It was easy to hear and understand the instructor.

Almost always 5 4 3 2 1 Almost never

Data analysis

Processing time (in milliseconds) was analyzed and compared across groups using the R (R Core Team, 2012) and the lme4 package (Bates, Maechler, & Bolker, 2012). Four groups were derived from combinations of the teaching assistants' and the participants' language backgrounds (i.e., native English speakers (English S) listening to NSTA, native English speakers (English S) listening to ITA, native Mandarin speakers (Mandarin S) listening to NSTA, and native Mandarin speakers (Mandarin S) listening to ITA). These four combinations were compared using a linear mixed-effects model. The combinations of speaker and listener language background were entered as the fixed effect. Participants' ID and trial number were entered as random effects.

Each question in the comprehension questionnaire was assigned a score of three or four. The total score of the comprehension questionnaire is 27. For example, in the question, "According to Ekman, how many universal emotions do human beings have and what are they?", participants got one point for every two emotions they wrote down. If they wrote down all seven emotions, they got four points. A one-way between-subject ANOVA was conducted to compare the comprehension questionnaire scores among different combinations of speakers' and listeners' language background.

The attitudinal questionnaire used 5-point Likert scales to obtain speaker ratings. The total scores were analyzed based on a one-way between-subjects ANOVA.

RESULTS

Processing

Compared to the native English listeners listening to the NSTA, the native English listeners listening to the ITA had significantly longer reaction time, the native Mandarin listeners listening to the NSTA also had significantly longer reaction time. Compared to the English listeners listening to the ITA, the native Mandarin listeners listening to the ITA had significant shorter reaction time. However, there is no significant difference between the reaction time of the English listeners listening to the NSTA and the Mandarin listeners listening to the ITA (see Table 2).

Table 2

Comparison of processing time across different combination groups.

Contrast	Estimate	SE	df	t-value	p-value
English-NSTA vs. English-ITA	-91.62	18.05	1320.31	-5.08	<0.0001***
English-NSTA vs. Mandarin-NSTA	-62.81	16.97	1290.38	-3.70	0.0013*
English-NSTA vs. Mandarin-ITA	-25.78	16.42	1328.59	-1.57	0.39
English-ITA vs. Mandarin-NSTA	28.81	17.69	1325.09	1.63	0.36
English-ITA vs. Mandarin-ITA	65.83	17.14	1290.35	3.84	0.0007**
Mandarin-NSTA vs. Mandarin-ITA	37.02	15.32	1329.06	2.42	0.07

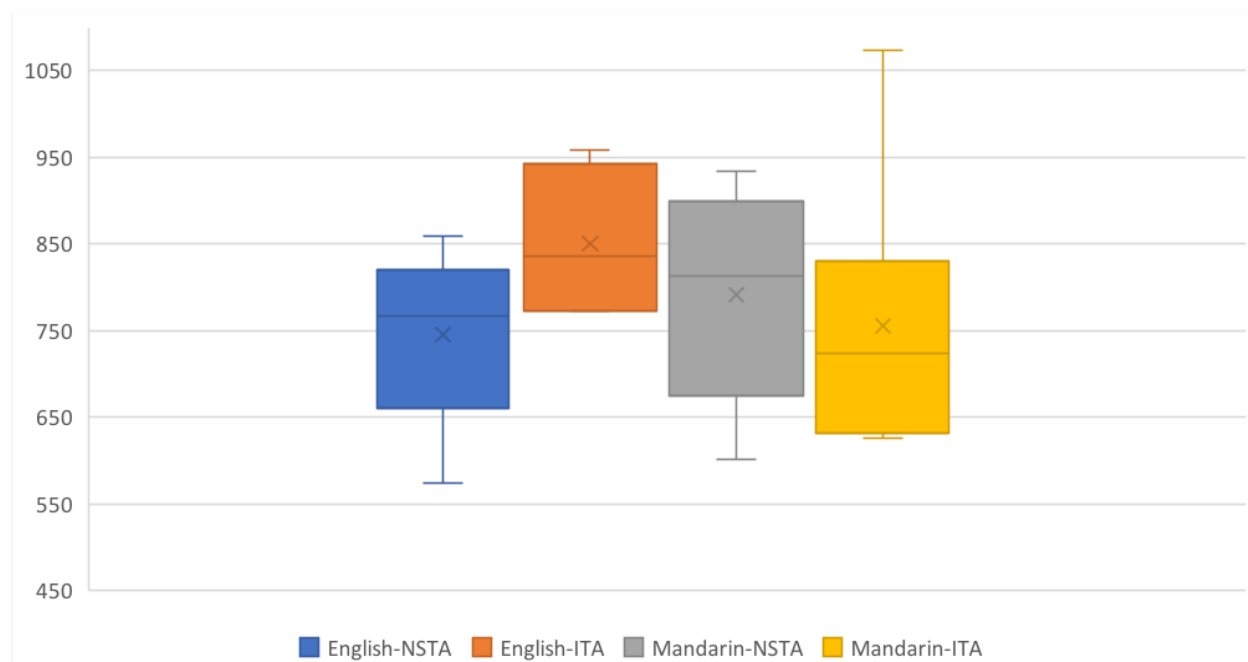


Figure 1. Processing time in the dual-task paradigm.

Comprehension

A one-way between-subjects ANOVA was conducted to compare the effect of listeners' language background on comprehension scores. The results show there was not a significant effect of listeners' language background on comprehension scores [$F(3, 17) = 0.36, p = 0.78$]. The post hoc analysis revealed that the statistical power for this analysis was 0.84. This result suggests that despite different processing difficulties, listeners' comprehension of the lecture content might not be significantly affected by the speakers' language background.

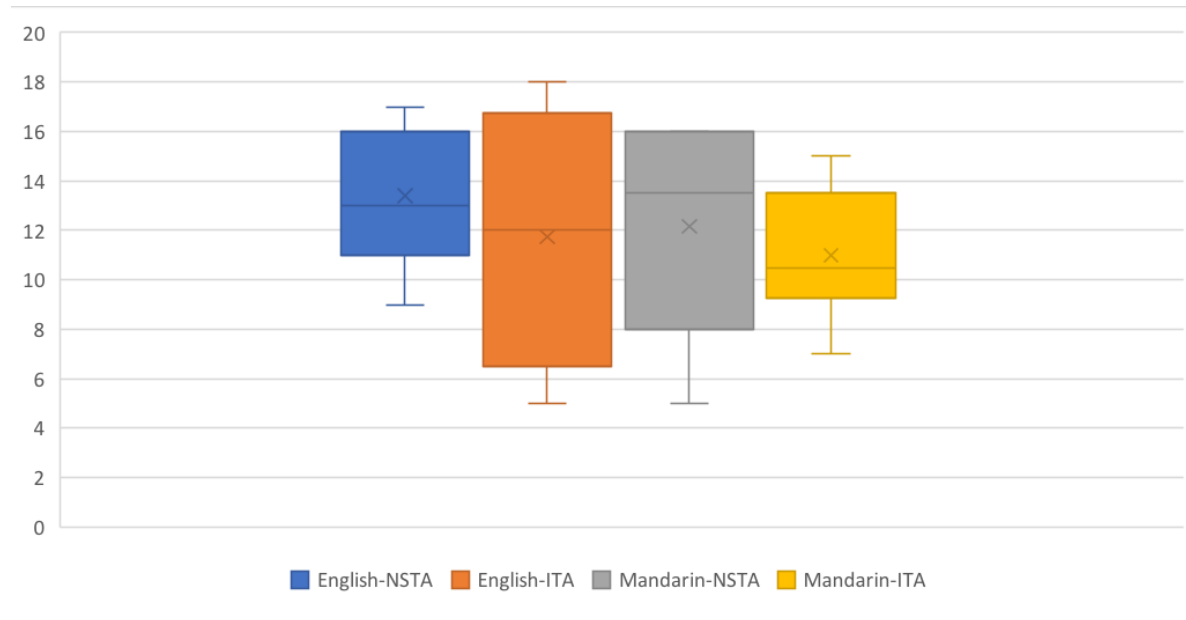


Figure 2. Comprehension scores of the four combination groups.

Attitude

A one-way between-subjects ANOVA was conducted to compare the effect of listeners' language background on attitude scores. The results show there was not a significant effect of listeners' language background on listeners' attitude towards the speakers [$F(3, 17) = 1.38, p = 0.28$]. The post hoc analysis revealed that the statistical power for this analysis was 0.99.

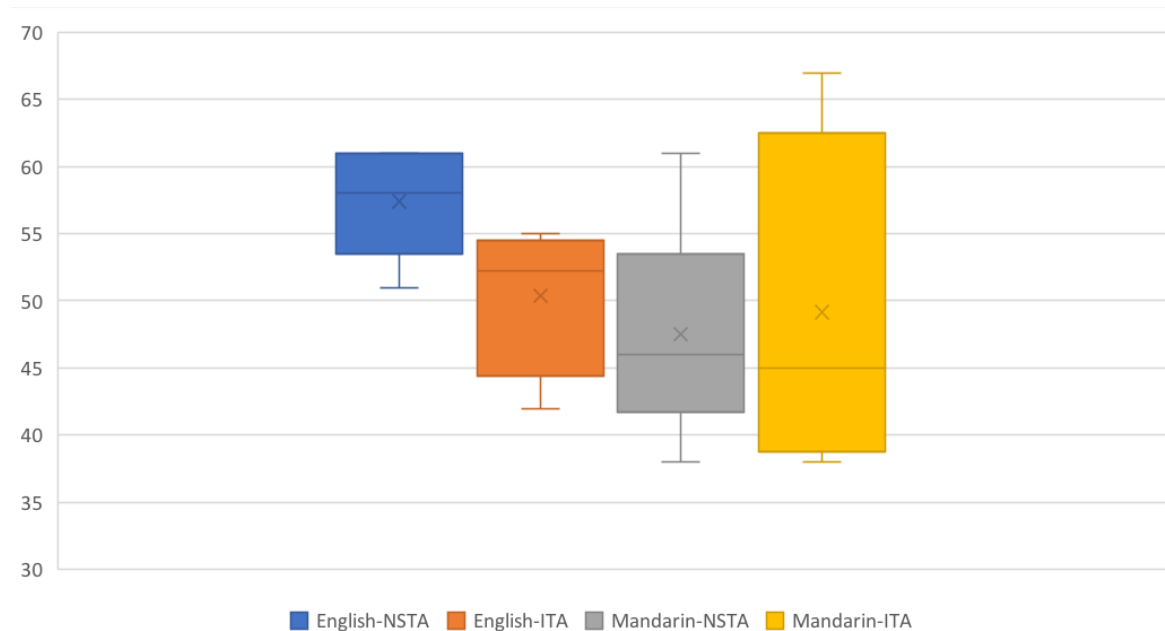


Figure 3. Attitudinal questionnaire scores of the four combination groups.

DISCUSSION

Using a dual-task paradigm, a comprehension questionnaire, and an attitudinal questionnaire, this study investigated whether native speaker teaching assistant (NSTA)'s and ITA's speech affects the processing, comprehension, and attitudes of listeners with different language backgrounds. The results show that the NSTA's and ITA's speech poses asymmetrical processing difficulties to listeners with different L1s. Specifically, it is easier for the native listeners to process the lecture given by the NSTA and it is easier for the non-native listeners than it is for the native listeners to process the lecture given by the ITA. These results confirmed prior researchers' findings on ITAs' pronunciation deficiency (Pickering, 2001) and the Interlanguage Speech Intelligibility Benefit (Bent & Bradlow, 2003) and extended them to the discourse level using a classroom lecture. However, the "mismatch interlanguage speech intelligibility benefit", which suggests that it is easier for non-native listeners to process the speech of a non-native speaker even when the listener and the speaker have different L1s, was not investigated in this study. It remains unsolved and should be investigated in future studies. Despite the difference in processing difficulties, there are no statistically significant differences in listeners' comprehension of the lecture or their attitudes towards the speakers. These results suggest that the processing difficulties may not necessarily lead to inferior learning outcomes or negative attitudes towards the speaker.

Some scholars attribute non-native speakers' deficiency in exploiting English pronunciation to the differences in the intonational system of English and the learners' L1. For example, Clennell (1997) argues that "the discourse/pragmatic functions of English prosody appear to be specific to the English language and are unfamiliar to most learners of English" (p.117). However, recent studies challenge this claim. For instance, scholars have found many similarities between English and Mandarin prosody, from physical correlations to pragmatic functions (Chen & Gussenhoven, 2008; Ip and Cutler, 2016; Liu, 2017; Ouyang & Kaiser, 2015). As Ip and Cutler (2016) concluded, "Information structure is a linguistic universal... all speakers have the option to convey this structure in the way they speak, and they may use prosody to do it" (p. 330). But if English and the learners' L1 share similar prosodic features and functions, why is it that non-native speakers do not transfer these functions to English? Multiple factors could lead to the lack of positive transfer.

It is possible that the non-native speakers are unaware of the functions and importance of prosody, both in their own language and in English. As Gilbert said, the importance of prosody is not "self-evident." In this case, raising learners' awareness of the importance and the pragmatic functions of English prosody might improve non-native speakers' use of English prosody, which could make it easier for native listeners to process the ITAs' speech.

It is also possible that even though both English speakers and Mandarin speakers use prosody in a similar manner and to a similar extent, English listeners rely more heavily on prosodic cues. This may explain why native listeners found it particularly challenging to process the ITA speech whereas non-native listeners, relying less heavily on prosodic cues, did not find it more challenging to process ITA speech.

Familiarity is another issue that could lead to the asymmetrical cognitive load imposed by the pronunciation and intonation difference of the NSTA's and ITA's speech. It has been found by

prior researchers (Gass & Varonis, 1984) that listeners' familiarity with the speakers' accent is a key variable in determining the intelligibility and comprehensibility of the speaker. Specifically, a non-native speaker's speech may be more intelligible to a listener who is familiar with the non-native speaker's accent. It is possible that the participants who are native speakers of English are less familiar with the Mandarin-speaking ITA's accent whereas the participants who are non-native speakers of English are familiar with both the Mandarin-speaking ITA's accent and the NSTA's accent. In other words, native listeners may have found it harder to process the ITA's speech because they are less familiar with Mandarin-accented English pronunciation.

ABOUT THE AUTHORS

Di Liu is a part-time lecturer at Boston University. His research focuses on cross-linguistic comparison of prosody, cognitive influence of L1 and L2 prosody, and technology-enhanced pronunciation teaching.

Marnie Reed is a professor of education at Boston University. Her research focuses on second language phonology. Her current areas of interest include the role of metacognition in cross-linguistic awareness of the pragmatic functions of English prosody.

REFERENCES

- Abercrombie, D. (1949). Teaching pronunciation. *ELT Journal*, 3(5), 113-122.
- 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.
- Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using Eigen and Eigenfaces. CRAN. R project.
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, 114(3), 1600-1610.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5(9-10), 341-345.
- Chen, Y., & Gussenhoven, C. (2008). Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 36(4), 724-746.
- Chun, D. (1988). Teaching intonation as part of communicative competence: Suggestions for the Classroom. *Die Unterrichtspraxis / Teaching German*, 21(1), 81-88.
- Clennell, C. (1997). Raising the pedagogic status of discourse intonation teaching. *ELT Journal*, 51(2), 117-25.

- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-397.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language learning*, 48(3), 393-410.
- Gass, S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language learning*, 34(1), 65-87.
- Gilbert, J. (2014). Myth 4: Intonation is hard to teach. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp.107-136). Ann Arbor, MI: University of Michigan Press.
- Liu, D. (2017). Mandarin speaker's intonational emphasis in English and Mandarin lectures. In M. O'Brien & J. Levis (Eds). *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Calgary, AB, August 2016 (pp. 85-96). Ames, IA: Iowa State University.
- Munro, M. J., Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97.
- Munro, M. J., Derwing, T. M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. M. Levis (Eds.), *The Handbook of English pronunciation* (pp. 377-396). West Sussex, UK: John Wiley & Sons, Inc.
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201-223.
- Hellermann, J. (2003). The interactive work of prosody in the IRF exchange: Teacher repetition in feedback moves. *Language in Society*, 32(1), 79-104.
- Ip, M. H. K., & Cutler, A. (2016). Cross-language data on five types of prosodic focus. *Proceedings of the Speech Prosody Conference 2016*, 330-334.
- 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(2), 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.
- R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

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.

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 University Press.