Reed, M. & Liu, D. (2020). Technology-enhanced L2 listening: Triangulating perception, production, and metalinguistic awareness. In O. Kang, S. Staples, K. Yaw, & K. Hirschi (Eds.), *Proceedings of the 11th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, Northern Arizona University, September 2019 (pp. 173–185). Ames, IA: Iowa State University.

TECHNOLOGY-ENHANCED L2 LISTENING: TRIANGULATING PERCEPTION, PRODUCTION, AND METALINGUISTIC AWARENESS

Marnie Reed, Boston University
Di Liu, Temple University

L2 listening poses challenges to learners and teachers alike. Learner challenges include parsing continuous speech, and identifying speaker intent (Vandergrift & Goh, 2012). Instructor challenges include overcoming what Field (2008) termed the "comprehension approach" whereby listening skills instruction consists of testing the accuracy of learners' comprehension. Learners' misattribution of their inability to recognize known words in connected speech to the speed of speech may be ameliorated by metalinguistic awareness of co-articulated divergence from citation form, coupled with technology-enhanced practice. Learners' resistance to adopting English prosody (Gilbert, 2014), save for imitation for mocking purposes, may be ameliorated by metalinguistic awareness of pragmatic and discourse functions of intonation plus practice of requisite pitch changes. A framework centered on mutual intelligibility and triangulating speech perception, production, and metalinguistic awareness is proposed. Based on research findings supporting the effectiveness of technologies such as high variability input (Thomson, 2012) and speech visualization (Levis & Pickering, 2004), technological tools including *English Accent Coach, POSE-test, YouGlish* and *Praat* are recommended to facilitate learner awareness-raising and teaching of segmental, suprasegmental, and connected speech features. We share the potential of technology-enhanced instruction to improve L2 listening skills.

INTRODUCTION

Listening has been described as the skill most widely used in daily life (Morley, 2001), the skill that develops faster than the other language skills (Rost, 2001), and the skill that can facilitate development of the other language skills (Oxford, 1990). By Nunan's (1998) estimate, listening consumes 50% of the time spent functioning in a foreign language. Nevertheless, in the words of Mendelsohn (1994), the Canadian linguist who advocated putting the 'L' in the Speaking, Pronunciation, Listening Interest Section (SPLIS) of the International TESOL Association, from the perspective of teaching, listening is the 'Cinderella' of the four skills:

> Despite a gradually increasing acceptance of the importance of listening comprehension for second language learners, the teaching of listening comprehension remains a somewhat neglected and poorly taught aspect of English in many ESL programs – the 'Cinderella' skill of ESL. (p. 9)

In the intervening years, as noted by Nation & Newton (2009), listening persists as "the least understood and most overlooked of the four skills (L, S, R, & W) in the language classroom" (p. 37). While materials writers and classroom teachers increasingly promote the use of authentic materials, Rost (1994) notes that learners struggle to understand the spoken language as native speakers actually use it (p. 142). Documented learner challenges include segmenting speech that

is continuous, i.e., "without robust or reliable [word] boundary signals" (Cutler, 2012, p. 170), and the ability to use prosodic information to "infer the speaker's intention" (Vandergrift & Goh, 2012, p. 25).

Instructor challenges may reflect a combination of the Communicative Language Teaching (CLT) framework and the influence of reading pedagogy that have resulted in test-oriented listening instruction. As characterized by Mendelsohn (2006), "much of what is traditionally mis-named *teaching* listening should in fact be called *testing* listening" (p. 75, italics in the original). The curriculum guidelines that teachers may turn to for direction seem actually to conflate reading and listening objectives: e.g., understand the main points and significant details. Further, based on a compilation and examination of textbooks used for Listening & Speaking courses in local community college and IEP classrooms, the pedagogic intervention adopted in the classroom seems to be to teach note-taking skills. However, the prerequisite underlying skills differ: reading requires ability to decode orthographic input, while listening requires ability to process aural input. The distinction between reading and listening becomes immediately clear for anyone who has worked with pre-literate Students from Limited or Interrupted Formal Education (SLIFE) populations. There is a recognized futility in handing out a written passage and expecting students to read and answer the comprehension questions. There seems not to be a corresponding recognition by teachers that prior to playing a listening passage, students must be taught how to process the spoken input.

Our framework takes a position on the 'which comes first: perception or production' question. Linebaugh & Roche (2015) demonstrated the efficacy of listening instruction that promotes production training of problematic sounds to facilitate second language perception. This supports a focus suggested by Casserly and Pisoni (2010) on shaping a speaker's own speech production to activate auditory feedback. We advocate extending the benefits of speech production training in informing and facilitating speech perception at the segmental level to addressing connected speech processes (CSPs) and suprasegmental features.

Our framework also addresses learner beliefs with respect to their challenges in reconstructing what was said (communicated content), and discerning what was meant by what was said (communicative intent). In describing the 'native listener' influence on second- or foreign-language listening, Cutler (2012) argues that speech processing is shaped by experience listening to the native language. One result, noted by Broersma and Cutler (2008), is lexical competition, characterized by substitution of known words for unrecognized words, accompanied by inability to suppress wrong choices even when not contextually supported. Unaware that words heard embedded in a continuous speech stream lose clear boundaries, listeners may misattribute their inability to recognize known words in connected speech to the speed of speech. Technology can increase metalinguistic awareness of the cues that are most effective for the target language, and promote flexibility in adjusting to its characteristics.

Regarding learners' attitudes about English prosody, Paunović and Savić (2008) assert that "Students often do not have a clear idea of why exactly 'the melody of speech' should be important for communication, and therefore seem to lack the motivation to master it…" (p. 72). They further assert that "teachers do not seem to be theoretically or practically well-equipped to explain and illustrate its significance" (Paunović & Savić, p. 73). Mindful of Levis (1999), who cautioned

against the historic textbook and instructional (mis)focus of intonation on conveying speakers' attitudes and emotions, we avoid emphasis on sarcasm and the various affective functions of intonation. Instead, our concern is with the discourse and pragmatic functions of intonation. Unaware of these, learners may "think intonation is merely decorative" (Gilbert, 2014, p. 125).

While learners may opt not to adopt intonation in their own speech, nevertheless, they need to be able to detect a marked pitch range and to attribute speaker intent to the signal it sends, such as emphasis, contrastive or corrective information, or implication of unstated information not retrievable from the utterance alone. Consider a common classroom exchange, reported by Reed and Michaud (2015), wherein a student requests permission to turn in an assignment late. Attending only to the (affirmative) words, "You *can*." But not the (negative) message conveyed by intonation (signaled here by italics), the student responded with relief, "Okay, thanks" (p. 463). Here, too, tech tools promote awareness raising and practice, for example, with what Wells (2006) terms the implicational fall-rise pitch contour.

THE TRIANGULATING FRAMEWORK

A framework centered on mutual intelligibility, and triangulating speech perception, production, and metalinguistic awareness is proposed (see Figure 1).

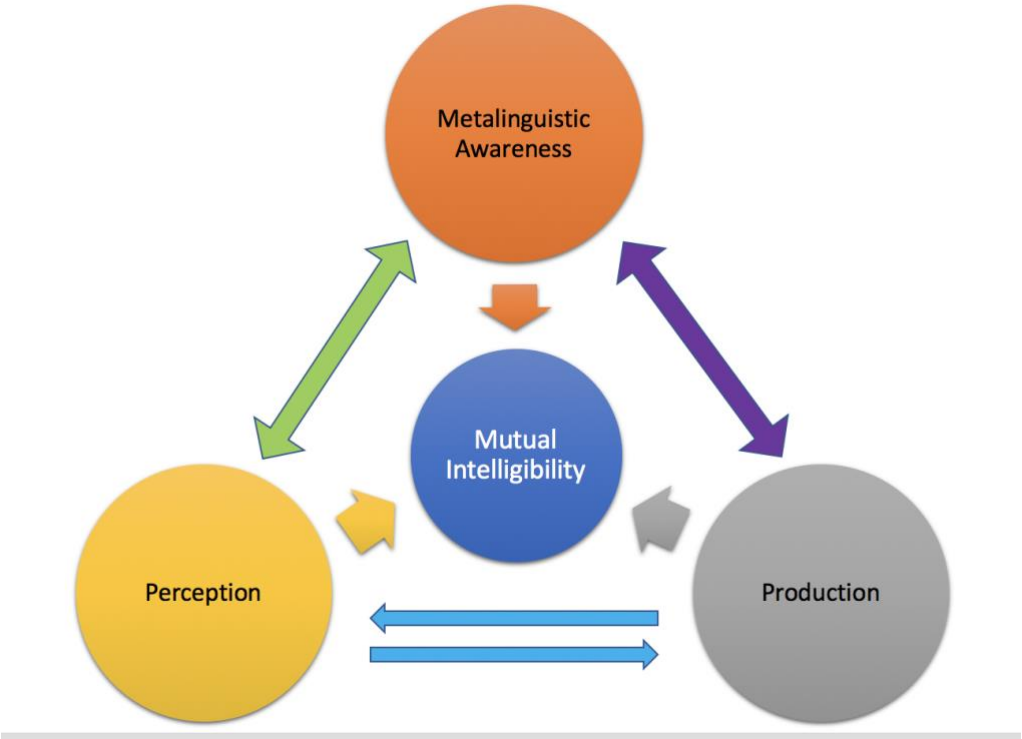


Figure 1. Triangulating perception, production, & metalinguistic awareness

Mutual Intelligibility

The core of our proposed framework is mutual intelligibility. According to Munro and Derwing (1999), "Intelligibility may be broadly defined as the extent to which a speaker's message is actually understood by a listener" (p. 289). This entails a role for both speaker and listener. With respect to the speaker, "It is almost universally acknowledged that the goal for adult L2 learners in regard to their own pronunciation is intelligibility" (Wagner & Toth, 2016, p. 87). With respect to the listener, "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" (Munro & Derwing, 2015, p. 388). Specifically, listeners' background and characteristics such as familiarity with accent and content, attitudes, language proficiency, and linguistic awareness may substantially influence their perception of intelligibility (Ginther & Yan, 2017). Mutual intelligibility requires learners "both to make their speech understood and to understand the speech of others" (Levis, 2018, p.1). With respect to the two challenges we've been discussing, Levis (2017) claims that this broad definition implies at least two different types of understanding: successfully identifying words and understanding a speaker's intended meaning.

Metalinguistic Awareness

Formerly construed as language awareness (Hawkins, 1987), metalinguistic awareness is more recently viewed as "the ability to reflect on language as a symbolic system in its own right" (Roberts, 2011, p. 45). As applied to connected speech processes, this entails awareness that words are contracted and linked; sounds are reduced and altered, etc. As applied to prosody, this entails awareness that stressed syllables or words are longer in duration, higher in pitch, and greater in intensity. In addition to awareness of intonation's role in signaling grammatical relations or conveying speaker attitude and emotions, learners need to be aware that intonation also "has the power to reinforce, mitigate, or even undermine the words spoken" (Wichmann, 2005, p. 229).

Perception and Production Loop

Granted the received wisdom that perception precedes production in second language speech learning, a number of studies have successfully challenged this precedence relationship. For example, results of a study reported by Linebaugh and Roche (2015), which examined perceptually assimilated sounds, demonstrated that "learners are able to utilize auditory feedback from their own speech to shape, adjust, or define phonetic categories in the second language, and those more accurate phonetic categories lead to improved perceptual ability" (p. A-9). The success of production training in facilitating perception at the segmental level illustrates the interrelated nature of speaking and listening. As noted by Wagner (2014), a main source of spoken input in the language classroom is "spoken language the learners themselves produce and self-monitor" (p. 289). This suggests that "the route to successful listening comprehension is through auditory feedback wherein the learner's own increasingly target-like speech production facilitates and reinforces perception" (Reed & Michaud, 2011. P. 95). We illustrate this relationship in Figure 2.

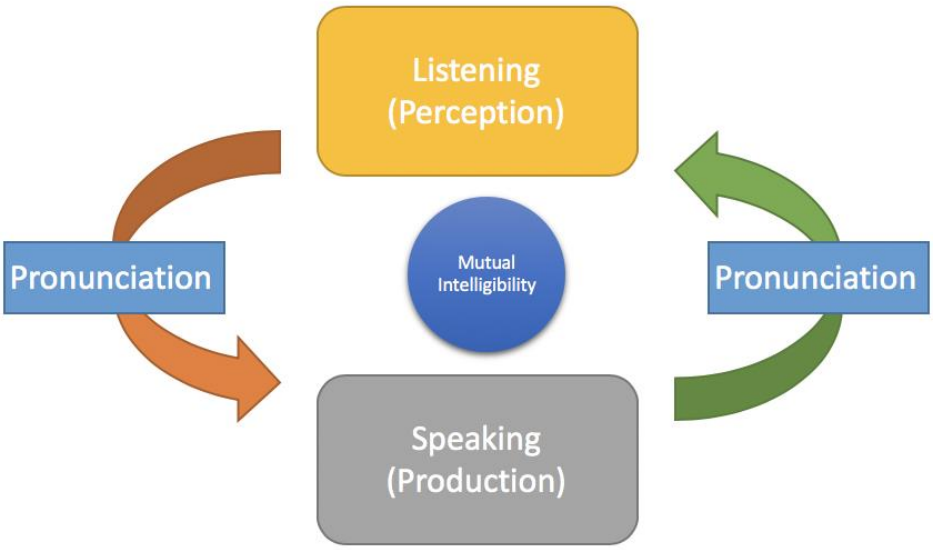


Figure 2. Speech perception and production loop

It is important to note that both the perception and production of speech are interwoven with metalinguistic awareness. Regarding perception, researchers have identified a significant positive relationship between learners' metacognitive awareness and listening performance (Goh & Hu, 2014). Regarding production, research shows that metalinguistic awareness enhancement can largely facilitate the production of some pronunciation features, such as sentence stress (Liu, 2018).

TECHNOLOGIES

Recent years have witnessed a rapid growth of interest and application of computer-assisted pronunciation teaching (Levis, 2007; Levis & Pickering, 2004; Thomson, 2011, 2012). However, despite progress in the development of technologies, there remain some issues.

Hincks (2015) stated that:

> in general, computers lend themselves most naturally to the kind of training advocated by audiolingual theorists: drills, repetitions, and mimicry. The theories of the communicative approach to language learning are harder to put into practice. Further advances in artificial intelligence are necessary before computers can offer an environment that can be said to truly either "communicate" or "negotiate" with a learner… (p. 506).

We generally agree with this insight, and acknowledge that no technology can yet serve as an interlocutor for language learners. However, a systematic integration of technologies may help teachers create a communicative environment which could enhance listening teaching by unifying speech perception, production, and metalinguistic awareness. This is achievable at multiple levels.

English Accent Coach (https://www.englishaccentcoach.com/) is a website that tests learners' perception of individual segments (i.e., consonants and vowels). It uses high-variability input, which trains learners to "perceive sounds produced by multiple talkers in multiple phonetic contexts and results in significantly greater improvement than training that relies on a single talker or single context" (Thomson, 2012, p.1235). English Accent Coach can generate a report card for each learner pointing out the sounds that need more work (see Figure 3). Using this report card, both teachers and students may have a better understanding of the learner's specific perception difficulties. Designed teaching activities targeting individual learners' perception difficulties, therefore, may be more effective than an aggregate teaching approach.

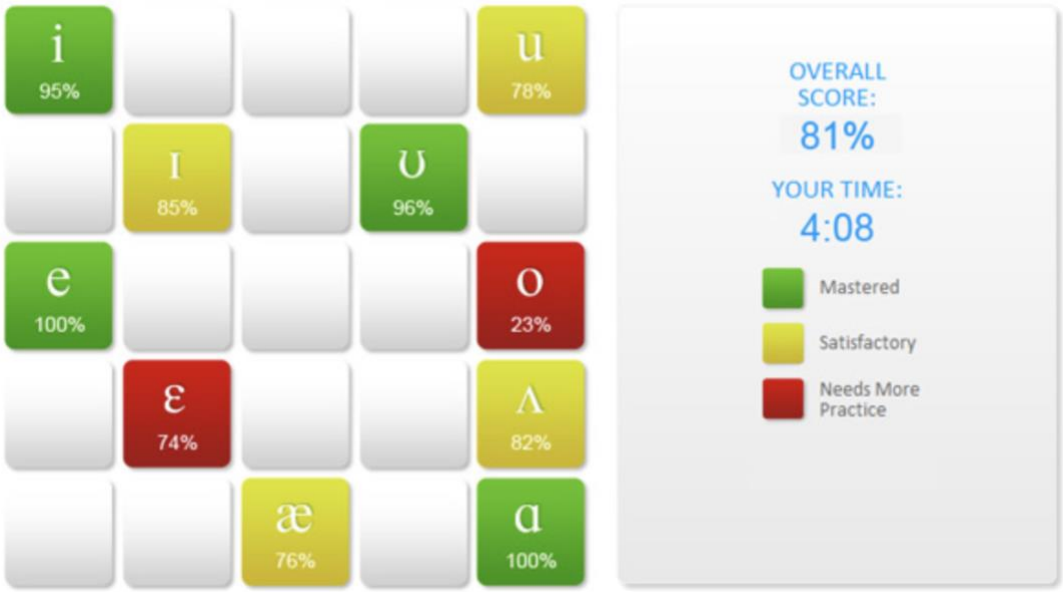


Figure 3. English Accent Coach report card

The scope of *English Accent Coach* is segmental features. For a more comprehensive treatment that includes suprasegmental features such as intonation and stress, teachers are advised to combine *English Accent Coach* with the *Perception of Spoken English (POSE) Test* (https://posetest.com/), which includes sections that ask students to associate certain intonation patterns with pragmatic functions. For example, in the "Sentence Stress 1" section of the website, learners will be asked to listen to a sentence and choose between two written sentences with the same words but different intonation (see Figure 4). The website can serve as a tool for both perception assessment and metalinguistic awareness enhancement.

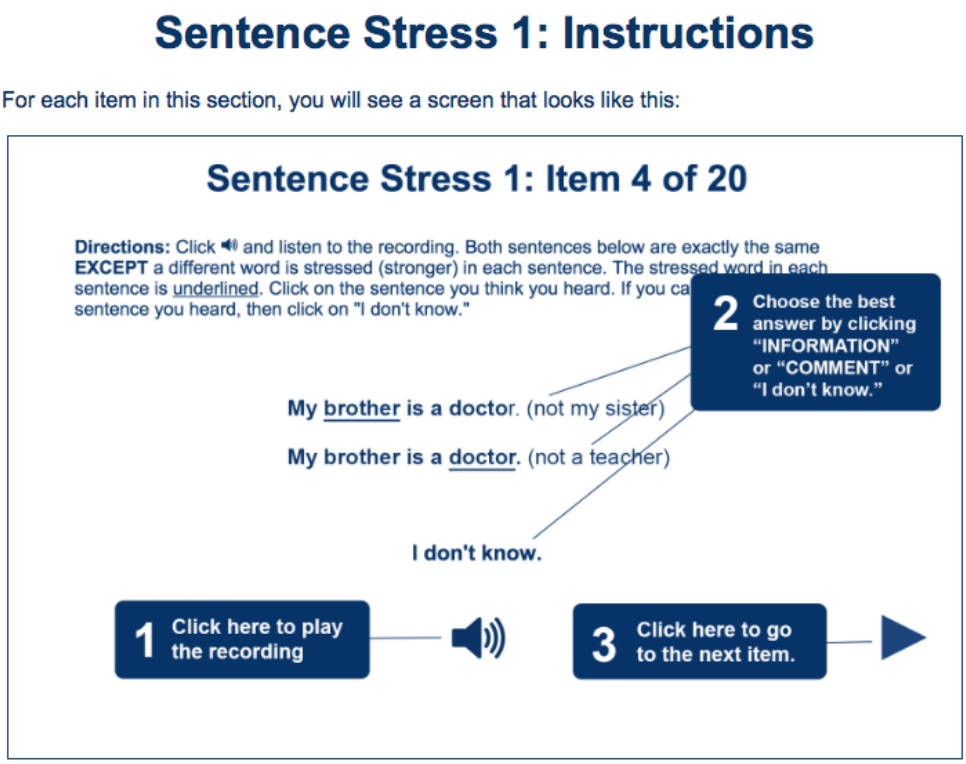


Figure 4. Perception of Spoken English (POSE) test demo

Technology Enhanced Listening Teaching

To address the listening challenges identified above, we present samples of the technologies integrated into the triangulated framework. Among the topics that Alameen and Levis (2015) recommend addressing in the second language classroom are connected speech processes (CSPs) such as linking, deletion, reduction, and so on. This is because "words are not produced in an isolated fashion but rather have a tendency to "run together" (Celce-Murcia et al., 2010, p. 163), often rendering their pronunciation unrecognizable from their citation form.

YouGlish is a website that allows learners to search videos with keywords. Using authentic materials at YouGlish (https://youglish.com/), teachers may design teaching activities targeting CSPs relevant to both broad (Is he busy? "izzybizzy" [IzibIzi]) and field-specific "the judge'll make a ruling" contexts.

A case in point, included in all textbook series and curricula, is the future with 'will'. Students are taught its function as a modal auxiliary, its co-occurrence with the unconjugated form of the main verb, constraints on its various uses, but not necessarily its pronunciation when contracted. Teachers attentive to students' pronunciation may notice that, when reading aloud sentences with contractions, students will typically avoid the contracted form. This became apparent during a class observation recently when the instructor returned to a reading lesson by asking a student to resume the reading passage they had been working on earlier at the section that began, "It'll". Try as he did, the increasingly distressed student, even with the whispered plea to a classmate for help,

could not find “It’ll” in the passage. Students who do not contract ‘will’ in their own speech will likely not recognize it when listening, especially accompanied, as in this case, by the flapped /t/ allophone [ɾ] in the word ‘it’ followed by the contraction.

YouGlish affords opportunities to hear words or phrases in continuous speech. Entering “going to” will generate samples with/without sandhi variation. Students can be asked to listen to a set number of video clips and calculate the percentage with reduced “gonna”. As for “it’ll”, with over 24,00 thousand “it’ll” videos, students are exposed to authentic, unscripted materials, receive high variability input, and learn with enhanced metalinguistic awareness (see Figure 5).

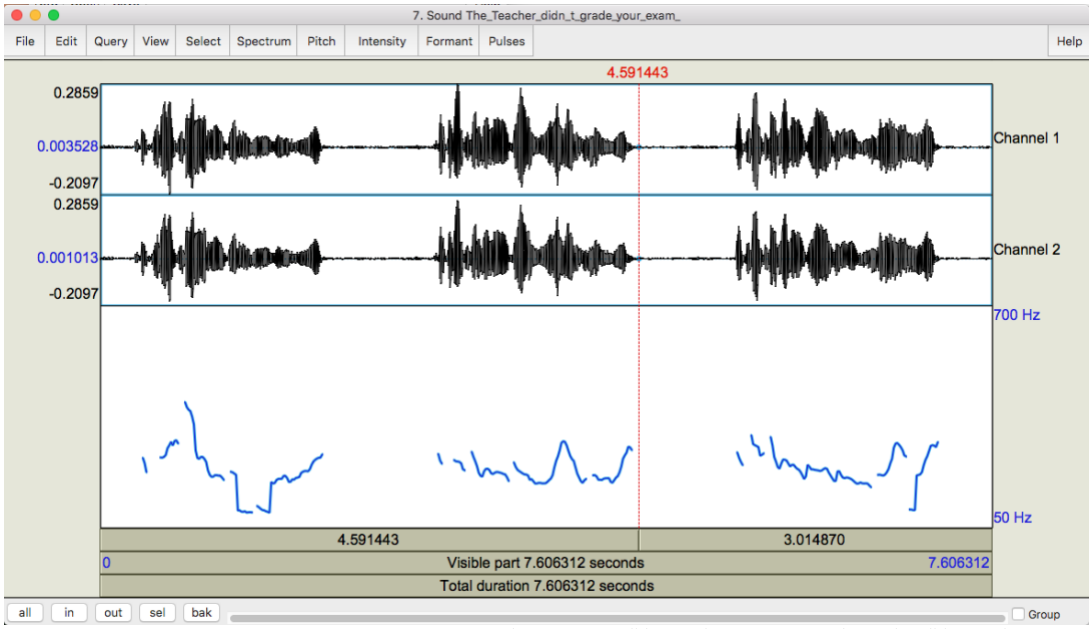
Figure 5. Searching “it will” at YouGlish.com

The importance of suprasegmental features has been widely acknowledged (Anderson-Hsieh et al., 1992; Celce-Murcia et al., 2010). For example, Derwing and Rossiter (2003) stated that “we

do not advocate eliminating segment-based instruction altogether, but, if the goal of pronunciation teaching is to help students become more understandable, then this study suggests that it should include a stronger emphasis on prosody [suprasegmental features]" (p. 14). Among the suprasegmental features, sentence stress (also referred to as focus, nuclear stress, primary stress, and prominence) is a high-value pronunciation feature crucial to comprehensibility (Levis & Levis, 2018). Hahn's (2004) study demonstrated that misplaced or missing primary stress in English discourse adversely impacted native English speakers' processing, comprehension, and evaluations of an International Teaching Assistant's speech. Even in speech among non-native speakers of English, Jenkins (2000) included emphatic and contrastive stress as crucial elements in her Lingua Franca Core.

Students may demonstrate, for example in coached laboratory practice, the ability to produce marked intonation contours and contrastive stress. However, as Gilbert (2014) observes, in the absence of instruction that raises awareness of the functions of these pitch contours, learners "may walk out of the class without having accepted the system at all" (p. 125). In a study reported by Reed and Michaud (2015), advanced-level students who dismissed unmarked English intonation as "exaggerated" or "sing-songy" failed to attend to the signal of marked intonation (represented here by italics) in interpreting the sentence, "The *teacher* didn't grade your exams." In responding to the question, "Have the papers been graded?", 70% of students in one class and 100% of students in another class (N = 14 in each class) incorrectly responded "No," revealing inability to discern the implication signaled by the marked intonation.

As Levis and Pickering (2004) state, "effective intonation practice without computers is ultimately restricted to teachers and students with good micro-listening skills" (p. 517). Speech visualization using software like Praat (Boersma & Weenink, 2018) has the potential to address this issue by presenting pitch movement of the same sentence with different sentence stress conveying different meanings. As illustrated in Figure 6, enhancing the noticing, locating, and interpreting of contrastive stress and implicational pitch contours can raise students' metalinguistic awareness by illustrating an aspect that may be new to some students whose L1s only accomplish sentence focus syntactically. Metalinguistic awareness enhancement and the dynamic nature of the listening task are both addressed.



The TEACHER didn't grade your exams. The teacher didn't grade YOUR exams. The teacher didn't grade your EXAM.

Figure 6. Using Praat to teach sentence stress

CONCLUSION

The Framework we proposed here is a systematic view of listening, speaking, and pronunciation teaching from problems, to theories, then to solutions supported by technology. We encourage pre- and in-service teachers to integrate technology in a cohesive approach to enhancing second language listening.

ABOUT THE AUTHORS

Marnie Reed is a Professor of Education at Boston University. Her current research interests include the role of metacognition in the cross-linguistic awareness of the pragmatic functions of English prosody.

Di Liu is an Assistant Professor of Instruction at Temple University. His research interests include the cross-linguistic comparison of prosody, the cognitive aspects of L1 and L2 prosody, and technology-enhanced language learning.

REFERENCES

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*(4), 529-555.

Alameen, G., & Levis, J. (2015). Connected speech. In M. Reed and J. Levis (Eds.), *The handbook of English pronunciation* (pp. 159–174). West Sussex, UK: John Wiley & Sons, Inc.

Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [Computer program]. Version 6.0.37, retrieved 14 March 2018 from http://www.praat.org/

Broersma, M. & Cutler., A. (2008). Phantom word activation in L2. *System, 36,* 22-34.

Casserly, E. D., & Pisoni, D. B. (2010). Speech perception and production. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(5), 629-647.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation: A course book and reference guide* (2nd Ed.). Cambridge: Cambridge University Press.

Cutler, A. (2012). Native listening: The flexibility dimension. *Dutch Journal of Applied Linguistics, 1*(2), 169-187.

Derwing, T. M., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, *13*(1), 1-17.

Field, J. (2008). *Listening in the language classroom*. Cambridge: Cambridge University Press.

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.

Ginther, A., & Yan, X. (2017). Listeners and raters: Similarities and differences in evaluation of accented speech. In O. Kang, & A. Ginther (Eds.), *Assessment in second language pronunciation* (pp. 79-100). New York: Routledge.

Goh, C. C., & Hu, G. (2014). Exploring the relationship between metacognitive awareness and listening performance with questionnaire data. *Language Awareness*, *23*(3), 255-274.

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

Hawkins, E. W. (1987) *Awareness of language: An introduction.* Cambridge: Cambridge University Press.

Hincks, R. (2015). Technology and learning pronunciation. In M. Reed and J. Levis (Eds.), *The handbook of English pronunciation* (pp. 505–19). West Sussex, UK: John Wiley & Sons, Inc.

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford, England: Oxford University Press.

Levis, J. M. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly*, *33*(1), 37-63.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics*, *27,* 184-202.

Levis, J. (2017). *Guidelines for promoting intelligibility*. Iowa State University slide share.

Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation.* Cambridge: Cambridge University Press.

Levis, J. M., & Levis, G. M. (2018). Teaching high-value pronunciation features: Contrastive stress for intermediate learners. *The CATESOL Journal*, *30*(1), 139.

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

Linebaugh, G., & Roche, T. (2015). Evidence that L2 production training can enhance perception. *Journal of Academic Language & Learning, 9*(1), A1-A17.

Liu, D. (2018). *Leveraging metalinguistic awareness and L1 prosody in the learning of L2 prosody: The case of Mandarin speakers learning English sentence stress*. [Unpublished doctoral dissertation]. Boston University.

Mendelsohn, D. (1994). *Learning to listen: A strategy-based approach for second language learners*. San Diego: Domine Press.

Mendelsohn, D. (2006). Learning how to listen using learning strategies. In P. Jorden (Ed.), *Current trends in the development and teaching of the four language skills* (pp. 75-89). Berlin: Mouton de Gruyter.

Morley, J. (2001). Aural comprehension instruction: Principles and practices. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 69-85). Boston: Heinle and Heinle.

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

Munro, M. J., & Derwing, T. M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed and J. Levis (Eds.), *The handbook of English pronunciation* (pp. 377–69). West Sussex, UK: John Wiley & Sons, Inc.

Nation, P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. New York: Routledge.

Nunan, D. (1998). Approaches to teaching listening in the language classroom. *Proceedings of the 1997 Korea TESOL Conference* (pp. 1-10). Taejon, Korea: KOTESOL.

Oxford, R. (1990). *Language learning strategies: What every teacher should know.* New York: Newbury House.

Paunović, T., & Savić, M. (2008). Discourse intonation— Making it work. In S. Komar & U. Mozetić (Eds.), *As you write it: Issues in literature, language, and translation in the context of Europe in the 21st century, V*(1-2), 57-75.

Reed, M., & Michaud, C. (2011). An integrated approach to pronunciation: Listening comprehension and intelligibility in theory and practice. In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference, Sept. 2010.* (pp. 95-104), Ames, IA: Iowa State University.

Reed, M., & Michaud, C. (2015). Intonation in research and practice: The importance of metacognition. In M., Reed & J. Levis (Eds.), *The handbook of English pronunciation*. (pp. 454-470) West Sussex: UK: John Wiley & Sons.

Roberts, P. (2011) *The role of metalinguistic awareness in the effective teaching of foreign languages.* Bern, Switzerland: Peter Lang.

Rost, M. (1994). *Introducing Listening*. London: Penguin.

Rost, M. (2001). Listening. In R. Carter and D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages* (pp. 7-13). Cambridge: Cambridge University Press.

Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal, 28*(3), 744-765.

Thomson, R. I. (2012). Improving L2 listeners' perception of English vowels: A computer-mediated approach. *Language Learning, 62*(4), 1231-1258.

Vandergrift, L. & Goh, C. (2012) *Teaching and learning second language listening: Metacognition in action.* NY: Routledge.

Wagner, E. (2014). Using unscripted spoken texts in the teaching of second language listening. *TESOL Journal*, 5(2), pp. 288-311.

Wagner, E. & Toth, P.D. (2016). The role of pronunciation in the assessment of second language listening ability. In T. Isaacs & P. Trofimovich (Eds.), *Second language pronunciation assessment: Interdisciplinary perspectives* (pp. 72-92). Bristol: Multilingual Matters.

Wells, J. (2006). *English intonation: An introduction*. Cambridge: Cambridge University Press.

Wichmann, A. (2005). The role of intonation in the expression of attitudinal meaning. *English Language and Linguistics, 9*(2), 229-253.