

INVITED REVIEW

Speech rhythm in English and applications to second language teaching

Donna Erickson^{1,2,*}¹*Sophia University, 7-1 Kioi-cho, Chiyoda-ku, Tokyo, 102-8554 Japan*²*Showa University of Music,**1-11-1 Kami-asao, Asao-ku, Kawasaki, 215-8558 Japan***Keywords:** Speech rhythm, Sentence stress, English, Second language applications, Acoustics, Articulation**PACS number:** 43.70.Kv, 43.70.Fq [doi:10.1250/ast.34.153]

1. INTRODUCTION

Verbal communication in a second language (L2) involves, among many things, skills in vocabulary, syntax, and pronunciation. Miscommunication can occur due to a slip in any one of these skills. Pronunciation errors in a second language are usually considered to be segmental (consonant and vowel) mispronunciations. A classic example is about Japanese speakers of English confusing /r-l/. For instance, the English driver's test examiner said "Turn at the next light" and the Japanese person turned at the next "right," and failed the test. Production and perception of speech sounds in L2 segments (consonants and vowels) are important for successful communication. However, prosody of a foreign language is also extremely important, yet something that has received relatively little attention both in research and second language training.

Prosody includes sentence intonation (the changes in fundamental frequency (F_0) which constitute the melody of an utterance), and sentence rhythm (changes in the stress, or accent, patterns in an utterance), e.g., [1]. Intonation and rhythm are different, although "the temporal correlation between intonational and rhythmic cues might actually facilitate the processing of speech rhythm" ([2], p. 85). Note that phrasing can often be at the discretion of the speaker, and choices in phrasing affect both sentence rhythm and sentence intonation. In addition, voice quality changes are part of prosody. The term prosody is the catch-all term for whatever is not traditionally expressed in the written form of the language. According to some researchers, prosody is at least as important, or more important than accurate segmental pronunciation [3].

Rhythm, as an ingredient of prosody, is also important, and a source of frequent miscommunication. A classic example is the 3-syllable English word, "McDonald's,"

with lexical stress on the second syllable, but spoken by many naive speakers of Japanese as an English loanword, as a 6-syllable word, "makudonarudo," with a high pitch on the third mora from the end of the word. The position of stress and the number of syllables have changed, thus changing the rhythm; listeners who have no familiarity with Japanese cannot understand the meaning. Minematsu *et al.* [4], examining transcriptions of sentences by American listeners of foreign accented Japanese English (17,416 responses), reports "very low intelligibility of Japanese English by Americans with little exposure to Japanese English," p. 1484. A study by Menezes *et al.* [5] recorded a Japanese bilingual speaker producing 10 true sentences (e.g., "Autumn is one of the four seasons") and 10 false sentences (e.g., "When it is hot, the sea becomes icy"), spoken with American English rhythm, and with Japanese English rhythm. The results showed that American college student listeners were able to correctly judge the veracity of the sentences spoken with native-like rhythm better than those with Japanese English rhythm. Rhythmic productions optimize perception by the listener, making speech easier to understand, e.g., [6].

To study the effect on comprehension, only due to wrong rhythm and ignoring segmental mistakes, is difficult for a number of reasons, including the fact that pronunciation of segments is integrally connected with rhythm [7]. Infants use rhythm as a sort of bootstrapping cue toward acquiring their native speech [2,7] around 6 months during the babbling stage [8]. First language rhythm is learned early (perhaps even in utero), learned well and difficult to change later in life.

2. WHAT IS RHYTHM?

Intuitively, we know that language has rhythm. But what is rhythm? Rhythm can be defined as any regular recurring motion, or movement marked by a regulated succession of strong and weak elements, a regular

*e-mail: EricksonDonna2000@gmail.com

recurrence or pattern in time referring to cyclical natural phenomena. A phonetic definition, taken from Kohler [9], is that rhythm is recurring (acoustic/articulatory) patterns of prominence within speech chunks (sentences, phrases, sub-phrases, feet); that is, it is sentence stress (accent) and the organization/patternings thereof that make up rhythm.¹ For spoken English, there tend to be alternations of relatively strong and weak syllables, which are chunked into phrases. Prominence in syllables and grouping of syllables into phrases are determined by the metrical phonological rules of the speaker, so that each syllable theoretically has a numerical value of stress, as determined by counting the number of hierarchical layers that the syllable participates in a particular utterance [10–12]. Note that from the metrical viewpoint, there is no primary or secondary sentence level stress; each syllable has a stress level based on its being a member of the sentence hierarchy. This is shown more clearly in Fig. 1 (Sect. 5). The hypothesis in this paper is that the numerical stress levels are reflected in acoustic and articulatory measurements, as discussed in [13], and explained in more detail of Sect. 5 of this review paper.

3. SENTENCE STRESS AND RHYTHM

Rhythm is the recurring patternings/the organization of sentence stress/accent throughout the utterance.² When sentence stress rules are violated, sentence rhythm is disrupted. Some basic English sentence stress rules include (1) content words are generally more strongly stressed than function words, (2) the first content word is more strongly stressed than the second one, in a series of content words, or with compound words, (3) phrase final lengthening occurs; and (4) nuclear accented word (primary sentence stress)/contrastively emphasized word has the most stress in the utterance.

There are rules for assigning sentence stress, such as, only one phrasal stress per phrase. These stress levels are metrically (hierarchically) generated, along the lines proposed by e.g., [10–12].

4. ACOUSTIC CHARACTERISTICS OF SENTENCE STRESS AND RHYTHM

The acoustic characteristics of lexical stress are increases in fundamental frequency (F_0), intensity, duration, e.g., [15], and vowel quality/formant changes, e.g., [6]; generally, these characteristics also apply for sentence stress. For instance, Cooper *et al.* [16], reporting on the acoustic characteristics of contrastive stress in English, showed that focused (contrastively stressed) words

showed an increase in duration, that varied depending on the location of the contrastive stress in the utterance. Findings for position effects on sentence stress are also reported by e.g., [17]. Cooper *et al.* [16] reported increased F_0 for contrastive stress; however, for English it is possible to have low F_0 or high F_0 on stressed words in a sentence, e.g., [18]. Work by Mori *et al.* [19], looking at acoustic recordings of 20 American English speakers suggests, that depending on the phrasing and mixture of content and function words within a phrase, stress may be implemented phonetically by a subset of the above acoustic characteristics. For instance, in sentences where content and function words alternate (and consequently, the function words have reduced vowels), stress seems to be implemented by changes in duration and vowel quality; in sentences with a stress clash situation (with four successive full vowels), pitch and intensity tend to be used in addition. Moreover, individuals may vary as to which cues they mainly use for which types of sentence stress. Much work, however, remains to be done with regard to acoustic implementation of sentence stress in spoken English.

With regard to rhythm, an approach has been to measure changes in duration between stressed syllables, e.g., [20] or, syllable duration as a function of the total number of syllables in a unit, e.g., [21], or, the relative duration of vocalic/consonantal intervals within an utterance, also referred to as “rhythm metrics,” e.g., [22]. In addition, rhythm studies have examined changes in sonority, e.g., [17,23].

But rhythm is not simply changes in phonetic variables. It is the global temporal bracketing of the speech signal into chunks that have recurring phonetic characteristics over and above the syntactic and semantic organization, and interacting with it [6]. The hierarchical (metrical) organization of sentence stress, as implemented acoustically and articulatorily, contributes to the speaker’s sense of rhythm [13], and the phonetic measures recruited vary depending on the speaker, the phrasing, the types of words, etc. that are in the utterance.

5. ARTICULATORY CHARACTERISTICS OF SENTENCE STRESS AND RHYTHM

Rhythm involves “regular recurring motion, or movement.” A speaker’s utterance rhythm can be implemented by patterned changes in syllable articulation (jaw, tongue, lip), which in turn result in patterned changes of duration, intensity, formant values, etc. of the syllable. Over the past several decades, with the availability of tools such as X-ray microbeam [24] and electromagnetic articulography (EMA) [25], researchers have investigated aspects of articulation of sentence stress, including emphatic stress, nuclear stress, etc., and have reported increased articulatory activity with sentence stress/contrastively emphasized

¹Metrically assigned sentence stress/accent is different from lexically assigned stress (lexical stress/accent) and how these interact is not addressed in this review paper.

²See [14] for a description for the different uses of the term ‘stress.’

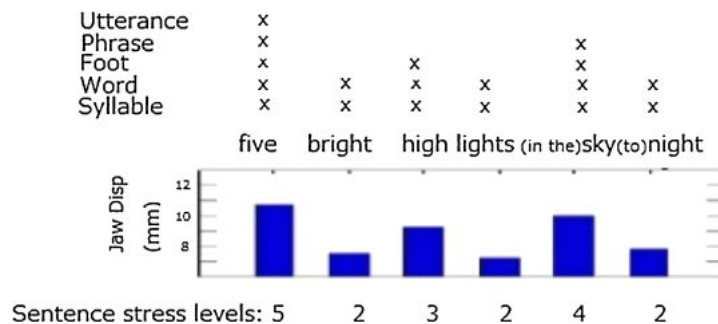


Fig. 1 Metrical structure (top part) and jaw displacement (X-ray microbeam data) (bottom part) for a single American speaker (average of 10 repetitions) (from [42]). The bottom row of numbers indicates the sentence stress levels, generated metrically based on the 5-level hierarchical organization of the sentence: syllable, word, foot, phrase, utterance. The sentence stress levels are assumed to be the sum of the x's for each level in which the syllable participates. Similar patterns were found for 3 other American English speakers, using EMA data [13].

words, particularly in connection with jaw opening, e.g., [26–32]. Also, see [33,34] for effects of prosodic structure on lip kinematics. Tongue movement is also more enhanced for sentence emphasis, e.g., [28,30]. Erickson [30] reported for emphasized words, the tongue moves more in the phonological direction of the vowel: for emphasized low vowels, it moves more down and back, and for emphasized high vowels, more up and forward. The acoustic result of this is a compression of the first two formant frequencies for the low vowel, and an expansion (diffusion) for the high front vowels.

The position of the syllable in the phrase/sentence affects the amount of jaw opening (see e.g., [29,35]), with the syllable/word at the head of a phrase generally having greater jaw displacement than that same syllable/word in other positions in the utterance, unless nuclear stress is assigned to a word in a later position, and then that word has the largest amount of jaw opening, e.g., [13]. Acoustic consequences of increased jaw opening have been reported to include increases in duration [32], intensity [36], and the first formant resonance frequency of the vocal tract (F_1) since the greater the jaw opening, the higher the F_1 , e.g., [13,32]. As for F_0 , Menezes *et al.* [18] showed that the jaw opens more for contrastively emphasized words, but that F_0 can be either high or low, depending on which accent pattern the speaker chooses to use.

Jaw displacement is constrained by height characteristics of the vowel nucleus (i.e., low vowels have more open jaw than high vowels); yet, studies by [37] for high and low vowels, and by [30,31] for high, mid and low vowels, have shown that regardless of the vowel height, increased prominence of a syllable in English results in increased jaw opening. For focused *vs.* non-focused Japanese vowels, increased jaw opening is also seen, but not to the same extent as for English [38].

The finding of increased jaw opening for syllables with larger sentence stress is in keeping with the C/D

(Converter/Distributor) Model of prosodic control of articulatory behavior, which purports that the larger the syllable prominence, the greater the articulatory movement, in particular, jaw opening [39]. Basically, the CD model is “a generative description of articulatory gesture organization for utterances, with an input made up of specifications for syllables by features, a parophonologically augmented metrical structure, and system parameters for utterance conditions” (p. 128). The skeletal rhythmic structure of the utterance is “represented as a syllable-boundary pulse train...” (p. 128); the magnitude of each syllable pulse is based as a first approximation on the amount of jaw opening for that syllable, and thus roughly represents the amount of prominence of each syllable.

Syllable onset labial consonants in English perhaps do not significantly affect the amount of jaw opening during the syllable nucleus, e.g., [40]; however, the vowel quality has a significant effect, with low vowels having more jaw opening than high vowels [41]. In order to understand articulation of sentence stress/rhythm, an algorithm for neutralizing vowel quality is also needed [41].

An approach to analyzing articulation of sentence stress/rhythm within a metrical framework is provided in [13]. This study examined four American English speakers kinematic (EMA and X-ray microbeam) patterns of jaw opening at the time of maximum jaw opening during the vocalic portions of the syllables, along with the corresponding F_1 values, in productions of a sentence containing all low vowels for the content words. Significant correlation of F_1 with jaw opening for three of the four speakers, as well as strong-weak jaw opening alternations corresponding to metrically generated syllable stress levels based on the sentence hierarchy for all four speakers, were found (See Fig. 1 below). The sentence hierarchy in Fig. 1 is composed of 5 levels: syllable, word, foot, phrase, utterance. Note there are two phrases in this utterance. F_0 and duration increased for the nuclear stressed syllable, but

no alternating strong-weak pattern was seen for F_0 and duration [42]. (No intensity measurements were made for this experiment.)

One interpretation of these results is that jaw deviation patterns reflect syllable magnitude, according to the definition of the C/D model. They affect the acoustic quality of the vowel (formants), and thus, the perception of patterns of sentence stress may be mediated mainly through patterns of changes in formant frequencies. In this view, implementation of rhythm (controlled alternations in sentence stress patterns) is primarily articulatory rather than phonatory, and the most consistent acoustic manifestation of rhythmic control may be vowel quality (formant) modulation. Changes in F_0 , duration and intensity also occur, with certain (not well understood) correlations between articulation (including jaw control) and phonation as well as the temporal organization control. Depending on the particular sentence stress pattern (i.e., rhythmic control), and the individual speaker and speaking style, these phonetic manifestations can vary, be mixed differently, and may or may not accompany alteration of the phonological metrical structure of the linguistic form used for the utterance, see e.g., [43]. More work on the rhythm and its articulatory manifestation, i.e., articulation of sentence stress pattern, is needed in order to corroborate or refute this hypothesis.

6. HOW TO CHANGE RHYTHM: APPLICATIONS TO SECOND LANGUAGE TEACHING

Rhythm of one's first language (L1), since it is learned as an infant, is difficult to change, but not impossible if one practices diligently with much repetition, similar to how one gains expertise in any performance skill, see e.g., [44]. According to Hallam [45], the more one practices, the better one gets, and the better one gets, the more one practices. The role of the second language teacher is to guide the learner to be perceptually sensitive to prosodic (including rhythmic) characteristics of the target language, and then to practice this repeatedly (in a group chorus works best), so as to be "cemented into his auditory memory as a template for future pronunciation in free production" ([44], p. 373).

Since English sentence stress is different from that of Japanese, a mora-timed, pitch accented language, stress is often perceived by Japanese learners in terms of increased F_0 instead of increased duration, intensity or changes in vowel quality, e.g., [19]. An English teacher thus needs to point out to Japanese students several aspects of English stress: Function words should be reduced in duration, e.g., [19]; sentence-initial function words should not be high-pitched, e.g., [46]; generally the last word in the phrase/sentence should be lengthened, e.g., [47]; and duration of

the first content word in an immediate sequence of two content words generally should be increased relative to that of the second word [19].

A more unified instruction to students might be to ask them to open their mouths more on the stressed words (and less on the unstressed words), thus resulting in appropriate increases in duration, intensity, as well as changes in vowel quality.

In order to help students learn to produce rhythm in a second language, biofeedback techniques using ultrasound, e.g., [48–50] are currently being explored.

After giving the learners guidance about rhythm production based on rhythm research, the teacher must then encourage them to practice and practice good rhythm [44]. Some methods of practicing include repeating after or with the teacher either individually or as a group [44], shadowing, e.g., [51], CALL training or other multimedia training, e.g., [52], or using speech recognition tools, including the latest speech recognition tools, such as "Siri."

The goal of the learner is to become sensitive to differences in L1–L2 rhythm, then, through practice, proceed to develop an internal sense of rhythm in the non-native language, and thus be able to communicate effectively in either language.

ACKNOWLEDGEMENTS

A grateful thanks to a number of people, including Osamu Fujimura, Klaus Kohler, Caroline Menezes, Yoko Mori, Nobuaki Minematsu, Olle Kjellin, Ian Wilson, Kanako Tomaru and especially to Takayuki Arai for the opportunity to write this review paper. This work was supported by the Japanese Ministry of Education, Science, Sport, & Culture, Grant-in-Aid for Scientific Research (C) #22520412. Note: The topic of spoken English rhythm and applications to teaching second language learners is larger than can be covered in this short article. My apologies to those researchers/studies who are not mentioned here.

REFERENCES

- [1] O. Fujimura, "Stress and tone revisited: Skeletal vs. melodic and lexical vs. phrasal," in *Proc. Symp. "Cross-Linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tone, and Descriptive Studies"*, S. Kaji, Ed. (Tokyo University of Foreign Studies, ILCAA, Tokyo, 2003), pp. 221–236.
- [2] F. Ramus, "Language discrimination by newborns: Teasing apart phonotactic, rhythmic, and intonational cues," *Annu. Rev. Lang. Acquis.*, **2**, 85–115 (2002).
- [3] J. B. Gilbert, *Teaching Pronunciation: Using the Prosody Pyramid* (Cambridge University Press, New York, 2008).
- [4] N. Minematsu, K. Okabe, K. Ogaki and K. Hirose, "Measurement of objective intelligibility of Japanese accented English: Using ERJ (English Read by Japanese) database," *Proc. Interspeech 2011*, pp. 1481–1484 (2011).
- [5] C. Menezes, D. Erickson, D. H. Whalen, T. Ito, Y. Mori and

- T. Hori, "Effect of non-native rhythm on native listeners comprehension of spoken utterances," *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 323–326 (2012).
- [6] K. J. Kohler, "Rhythm in speech and language: A new research paradigm," *Phonetica*, **66**, 20–45 (2009).
- [7] A. Cutler, "Segmentation problems, rhythmic solutions," *Lingua*, **92**, 81–104 (1994).
- [8] D. Crystal, "Prosodic development," in *Studies in First Language Development*, P. J. Fletcher and M. Garman, Eds. (Cambridge University Press, New York, 1986), pp. 174–197.
- [9] K. Kohler, "Whither speech rhythm research?," *Phonetica*, **66**, 5–14 (2009).
- [10] M. Liberman and A. Prince, "On stress and linguistic rhythm," *Linguist. Inq.*, **8**, 249–336 (1977).
- [11] E. O. Selkirk, *The Syllable* (Foris, Dordrecht, 1982).
- [12] B. Hayes, *Metrical Stress Theory: Principles and Case Studies* (The University of Chicago Press, Chicago, 1995).
- [13] D. Erickson, A. Suemitsu, Y. Shibuya and M. Tiede, "Metrical structure and production of English rhythm," *Phonetica*, **69**, 180–190 (2012).
- [14] K. Kohler, "The perception of prominence patterns," *Phonetica*, **65**, 257–269 (2008).
- [15] I. Lehiste, *Suprasegmentals* (MIT Press, Cambridge, Mass., 1970).
- [16] W. E. Cooper, S. J. Eady and P. R. Mueller, "Acoustical aspects of contrastive stress in question-answer contexts," *J. Acoust. Soc. Am.*, **77**, 2142–2156 (1985).
- [17] J. Edwards and M. Beckman, "Articulatory timing and the prosodic interpretation of syllable duration," *Phonetica*, **45**, 156–174 (1988).
- [18] C. Menezes, D. Erickson and O. Fujimura, "Contrastive emphasis: Comparison of pitch accents with syllable magnitudes," *Proc. Speech Prosody 2002*, Aix en Provence (2002).
- [19] Y. Mori, T. Hori and D. Erickson, "Acoustic correlates of sentence stress patterns for American English speakers and Japanese learners of English," *Phonetica* (in submission).
- [20] K. L. Pike, *The Intonation of American English* (University of Michigan Press, Ann Arbor, Mich., 1945).
- [21] I. Lehiste, "Rhythmic units and syntactic units in production and perception," *J. Acoust. Soc. Am.*, **54**, 1228–1234 (1973).
- [22] F. Ramus, M. Nespor and J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition*, **73**, 265–292 (1999).
- [23] N. Minematsu, S. Kobashikawa, K. Hirose and D. Erickson, "Acoustic modeling of sentence stress using differential features between syllables for English rhythm learning system development," *Proc. Int. Conf. Spoken Language Processing*, Vol. 2, pp. 365–368 (2002).
- [24] O. Fujimura, H. Ishida and S. Kiritani, "Computer-controlled radiography for observation of movements of articulatory and other human organs," *Comput. Biol. Med.*, **3**, 371–384 (1973).
- [25] T. Kaburagi and M. Honda, "Calibration methods of voltage-to-distance function for an electromagnetic articulometer (EMA) system," *J. Acoust. Soc. Am.*, **111**, 1414–1421 (1997).
- [26] M. Stone, "Evidence for a rhythm pattern in speech production: Observations of jaw movement," *J. Phonet.*, **9**, 109–120 (1981).
- [27] M. E. Beckman and J. Edwards, "Articulatory evidence for differentiating stress categories," in *Papers in Laboratory Phonology*, Vol. III, P. Keating, Ed. (Cambridge University Press, Cambridge, 1994), pp. 7–33.
- [28] K. de Jong, M. E. Beckman and J. Edwards, "The interplay between prosodic structure and coarticulation," *Lang. Speech*, **36**, 197–212 (1993).
- [29] D. Erickson, "Effects of contrastive emphasis on jaw opening," *Phonetica*, **55**, 147–169 (1998).
- [30] D. Erickson, "Articulation of extreme formant patterns for emphasized vowels," *Phonetica*, **59**, 134–149 (2002).
- [31] D. Erickson, "Some effects of prosody on articulation in American English," in *A New Century of Phonology and Phonological Theory, A Festschrift for Prof. Haraguchi*, T. Honma, M. Okazaki, T. Tabata and S. Tanaka, Eds. (Kaitakusha, Tokyo, 2003), pp. 473–491.
- [32] C. Menezes, "Changes in phrasing in semi-spontaneous emotional speech: Articulatory evidences," *J. Phonet. Soc. Jpn.*, **8**, 45–59 (2004).
- [33] T. Cho, "Manifestations of prosodic structure in articulatory variation: Evidence from lip kinematics in English," in *Laboratory Phonology 8: Varieties of Phonological Competence*, L. M. Goldstein, D. H. Whalen and C. T. Best, Eds. (Mouton de Gruyter, Berlin/New York, 2006), pp. 519–548.
- [34] M. Macchi, "Labial articulation patterns associated with segmental features and syllable structure in English," *Phonetica*, **45**, 109–121 (1988).
- [35] T. Cho and P. Keating, "Effects of initial position versus prominence in English," *J. Phonet.*, **37**, 466–485 (2009).
- [36] R. Schulman, "Articulatory dynamics of loud and normal speech," *J. Acoust. Soc. Am.*, **85**, 295–312 (1989).
- [37] J. Harrington, J. Fletcher and M. E. Beckman, "Manner and place conflicts in the articulation of Australian English," in *Papers in Laboratory Phonology*, Vol. 5, M. Broe and J. Pierrehumbert, Eds. (Cambridge University Press, Cambridge, 2000), pp. 40–51.
- [38] D. Erickson, M. Hashi and K. Maekawa, "Articulatory and acoustic correlates of prosodic contrasts: A comparative study of vowels in Japanese and English," *Proc. Spring Meet. Acoust. Soc. Jpn.*, 265–266 (2000).
- [39] O. Fujimura, "The C/D model and prosodic control of articulatory behavior," *Phonetica*, **57**, 128–138 (2000).
- [40] Y. Abe, I. Wilson and D. Erickson, "Video recordings of L1 and L2 jaw movement: Effect of syllable onset on jaw opening during syllable nucleus," *J. Acoust. Soc. Am.*, **132**, 2005 (2012).
- [41] J. C. Williams, D. Erickson, Y. Ozaki, A. Suemitsu, N. Minematsu and O. Fujimura, "Neutralizing differences in jaw displacement for English vowels," *Proc. Spring Meet. Acoust. Soc. Am.* (2013).
- [42] D. Erickson, S. Shibuya and A. Suemitsu, "Rhythm and emphasis in American English: Comparison of native and non-native speakers' productions," *Proc. Int. Semin. Speech Production, Montreal*, pp. 345–352 (2011).
- [43] W. E. Cooper and S. J. Eady, "Metrical phonology in speech recognition," *J. Mem. Lang.*, **25**, 369–384 (1986).
- [44] O. Kjellin, "Accent addition: Prosody and perception facilitates second language learning," in *Proc. LP'98*, O. Fujimura, B. D. Joseph and B. Palek, Eds. (The Karolinum Press, Prague, 1999), Vol. 2, pp. 373–398.
- [45] S. Hallam, "What do we know about practicing? Towards a model synthesising the research literature," in *Does Practice Make Perfect? Current Theory and Research on Instrumental Music Practice*, H. Jørgensen and A. C. Lehmann, Eds. (Norges musikkhøgskole, NMH-publikasjoner, Oslo, 1997), Vol. 1, pp. 179–231.
- [46] Y. Mori, "The initial high pitch in English sentences produced by Japanese speakers," *English Linguist.*, **22**, 23–55 (2005).
- [47] Y. Mori, "Phonetic realization of syntactic boundaries in English discourse produced by English vs. Japanese speakers," *J. Phonet. Soc. Jpn.*, **10**, 71–82 (2006).
- [48] B. Gick, B. Bernhardt, P. Bacsfalvi and I. Wilson, "Ultrasound imaging applications in second language acquisition," in

Phonology and Second Language Acquisition, J. G. H. Edwards and M. L. Zampini, Eds. (John Benjamins, Amsterdam, 2008), pp. 309–322.

- [49] J. S. Levitt and W. F. Katz, “Effects of EMA-based augmented visual feedback on the English speaker’s acquisition of the Japanese flap: A perceptual study,” *Proc. Interspeech 2010*, pp. 1862–1865 (2010).
- [50] H. Hatano, T. Kitamura, J. Nakamura, R. Hayashi and D. Erickson, “Comparison of tongue position during production of English non-stressed vowels uttered by native English/Japanese speakers,” *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 315–316 (2012) (in Japanese).
- [51] Y. Mori, “Shadowing with oral reading: Effects of combined training on the improvement of Japanese EFL learners’ prosody,” *Lang. Educ. Technol.*, **48**, 1–22 (2011).
- [52] D. Erickson, Y. Tanaka, K. Hirose and N. Minematsu, “Acquisition of L2 rhythm,” *Proc. Autumn Meet. Acoust. Soc. Jpn.*, pp. 241–244 (2011).