

Acquiring Rhythm: A Comparison of L1 and L2 Speakers of Canadian English and Japanese

Izabelle Grenon¹ and Laurence White²

¹University of Victoria and ²University of Bristol

Lacking knowledge of either language, one can readily distinguish Spanish and German speech, but distinguishing Italian and Spanish is more difficult. One reason that certain pairs of languages sound distinct is that they have different rhythms. The rhythmic contrast between, for example, Spanish and German arises in part from durational differences between stressed and unstressed vowels, and from the complexity of permissible syllable structures (Dauer, 1983). In German or English, for instance, the alternation between long, stressed vowels and short, reduced vowels is said to create a percept of a contrastive or “Morse code” rhythm (Lloyd James, 1940). Conversely, the much lower degree of stress-related lengthening, the relative attenuation of vowel reduction in unstressed syllables, and the lower frequency of complex consonant clusters in Spanish contribute to the impression of a more regular or “machine gun” rhythmic pattern, at least to speakers of English or German.

Based on these observations, a number of acoustic metrics for speech rhythm have been proposed (e.g. Dellwo & Wagner, 2003; Grabe & Low, 2002; Ramus, Nespore and Mehler, 1999; White & Mattys, 2007a, 2007b). Some of these measurements have been shown to be quite successful in predicting infants’ and adults’ perception of languages as similar or different, at least when speech is acoustically modified to focus on rhythmic information (Ramus, Dupoux & Mehler, 2003; Ramus et al. 1999; White, Mattys, Series, & Gage, 2007). Relatively little is known, however, about how the rhythm of one’s first language (L1) impacts on rhythmic production in a second language (L2), and about the utility of the acoustic metrics in quantifying this relationship.

A few studies have evaluated the influence of L1 on L2 production of rhythm using various acoustic metrics and methodologies. Low, Grabe, &

* A heartfelt thanks to Asuka Endo and Miwako Tateishi for their help with the Japanese materials and recordings, to Kyoko Kaneko for preparing the Japanese sentences, and to all our participants. Thank you also to John Archibald, Allison Benner, Darlene LaCharité, and Tae-Jin Yoon for their comments on earlier versions of this paper. This research has benefited from the support of the Social Sciences and Humanities Research Council of Canada (CGS award #767-2006-1176), the Fonds québécois de la recherche sur la société et la culture (FQRSC award #104021) granted to Izabelle Grenon, and a Leverhulme Trust (UK) grant to Sven Mattys (number F/00 182/BG).

Nolan (2000) examined the influence of L1 Chinese on L2 Singapore English; Gut (2003) evaluated the influence of L1 Chinese, English, French, Italian and Romanian on L2 German; Carter (2005) looked at Mexican Spanish influences on L2 American English; Lin & Wang (2005) studied the influence of L1 Chinese on L2 Canadian English; and White & Mattys (2007a) evaluated the L1-L2 interaction between Castilian Spanish and British English, and between Dutch and British English. Typically these studies – comparing languages that differ from each other in duration and distribution of stressed syllables, use of vowel reduction, and syllabic complexity – have shown rhythm scores for L2 speakers to be intermediate between those of native speakers of the L1 and L2.

Japanese differs from all of the languages mentioned above in important regards. First, Japanese does not have stress or vowel reduction, being a pitch-accent language in which the accented vowels are not generally lengthened (Akamatsu, 1997; Tsujimura, 1996; Vance, 1987). Second, as a quantity language, Japanese uses duration to distinguish both vowel and consonant phonemes: short and long vowels, and short and geminate consonants, are contrastive. In addition, Japanese syllable structures are fairly simple: there are no consonant clusters in onsets, and in coda position, only the first part of a geminate consonant or the nasalized obstruent /n/, often realized as a nasalized vowel, are allowed (Akamatsu, 1997).

By contrast, English employs stress and vowel reduction, exhibiting a phonetic contrast between long stressed vowels and short reduced vowels. English also allows a wide range of syllable structures, with complex consonant clusters both in onset and coda positions, particularly in stressed syllables.

Hence, both Japanese and English are predicted to exhibit a certain amount of variation in the duration of their vocalic and consonantal intervals, but this variation results from two essentially different phonological contrasts: long versus short phonemes in Japanese, and stressed versus unstressed syllables in English, together with a degree of phonemic vowel duration variation in many varieties of English (the “tense” vs “lax” vowel distinction, under which pairs of similar vowels differ in both quantity and duration, though the durational contrasts are much smaller than those in the stressed vs unstressed distinction).

The purpose of this study is to evaluate: (a) the interaction of L1 phonology with L2 rhythm production; and (b) the effectiveness of three acoustic metrics (%V, VarcoV and rPVI_C, discussed below) in identifying this interaction in the comparison of Canadian English and Japanese.

1. Acoustic metrics

A number of acoustic metrics have been proposed for the comparison of rhythm between and within languages (see White & Mattys 2007a, 2007b for a review and evaluation of these metrics). The metrics require segmentation of speech into vocalic and consonantal intervals, where such intervals are defined as all consecutive segments of the same type (vowel or consonant) irrespective of linguistic structure (e.g. ignoring syllable, morpheme, or word boundaries).

Of the metrics investigated here, %V, which was introduced by Ramus et al. (1999), measures the relative proportion of vocalic and consonantal intervals. It is calculated by summing the duration of the vocalic intervals within an utterance, and dividing this by the total duration of the utterance; the resulting ratio is expressed as a percentage.

VarcoV measures variation in the duration of vocalic intervals (it is analogous to VarcoC, proposed by Dellwo & Wagner, 2003). It is calculated by dividing the standard deviation of vocalic interval duration within an utterance by the mean duration of vocalic intervals, and multiplying by 100. The purpose of dividing by the mean is to normalize the metric for speech rate variation. White & Mattys (2007a, 2007b) have shown both %V and VarcoV to be robust to speech rate variation, and to be the most discriminative metrics for between- and within-language comparisons, as well as for the study of L1-L2 interaction between languages such as Spanish, Dutch, and English.

We also investigated the utility of rPVI_C, the raw Pairwise Variability Index for consonantal intervals suggested by Grabe and Low (2002). The rPVI_C is calculated by subtracting the duration of each consonantal interval from the duration of the preceding consonantal interval, summing the absolute values of these pairwise comparisons, and dividing by the total number of pairwise comparisons.

In contrast with VarcoV and %V, rPVI_C scores tend to be inversely correlated with speech rate (White & Mattys, 2007a). Grabe & Low (2002) suggest that normalization for consonantal metrics is undesirable, because it is likely to remove linguistically interesting variation, a view supported by the results of White & Mattys (2007a, 2007b). Despite potential problems with rPVI_C, we chose to investigate its utility in the comparison of English and Japanese. Given the different processes that underpin variation in the consonantal intervals of the two languages, such a metric seems essential to get a complete picture of the influence of L1 on L2 rhythm, keeping in mind the possible influence of speech rate on rPVI_C scores.

2. Experiment

2.1 Materials

Each speaker read five sentences in the target language, either English or Japanese (see Appendix A). The English sentences were those used in White & Mattys (2007a, 2007b), as modified from a larger set used in Nazzi, Bertoncini, & Mehler (1998). The Japanese sentences were constructed based on the English model by a native Japanese speaker who was given no instructions regarding rhythm.

The English sentences excluded the approximants /l/, /w/, /j/, and /r/ to increase the reliability of the segmentation procedure, given that the boundary between an approximant and a preceding or following vowel is difficult to ascertain. The exception to this was that the approximant /r/ was still present in coda position, and was included as part of the preceding vocalic interval. The

Japanese sentences were constructed along the same lines, excluding the sounds /j/ and /w/ and the Japanese flap in any position.

2.2 Participants

For the L1 groups, six native Canadian English speakers (Eng_{ENG}) and six native Japanese speakers (Ja_{JA}) were recorded for the experiment (three males and three females in each language condition).

All the Canadian English speakers were born and raised in southwestern Canada (i.e. southern British Columbia or Alberta), were between 19 and 35 years of age (mean 26), and spoke no languages other than English.

Four of the Japanese speakers were born and raised in the Tokyo area or surrounding regions and spoke the so-called standard Japanese dialect, while two of the female speakers were from Osaka. The Japanese participants had been living in Canada for one week to three months at the time of testing (mean 1.4 months), but did not speak English or any other L2s fluently. The Japanese speakers were between 22 and 29 years of age (mean 26).

For the L2 groups, we recorded six Canadian English speakers (Ja_{ENG}) from southwestern Canada speaking L2 Japanese and six native Japanese speakers (Eng_{JA}) from the Tokyo area speaking L2 English. The Canadian English speakers had all lived in Japan (mean 2 years, 4 months; five males, one female). These speakers ranged in age from 18 to 32 (mean 25). The Japanese speakers were all living in Canada (mean 2 years; two males, four females) and ranged in age from 21 and 33 (mean 27).

Speakers of the L2 groups were considered advanced learners, based on the last course level taken in the L2 (intermediate or advanced level) and on the amount of time spent in the country where the L2 was spoken.

The participants were all recruited and recorded in Victoria, British Columbia, and received a small honorarium for their participation. None reported any hearing or speech impairment, with the exception of one English-speaking participant (L1 group) who reported reduced hearing in her left ear, a condition which did not appear to affect her speech production.

2.3 Recording procedure

All speakers performed two tasks. The first, providing directions on a map, was not used for this study. For the second, participants recorded in English were asked to read the five experimental sentences, preceded and followed by five other sentences. Participants recorded in Japanese were asked to read the five experimental sentences preceded by five others.

Participants were given some time to practise reading the sentences silently before reading them aloud. They were instructed to read at a comfortable rate using their normal conversational voice, to avoid pauses within sentences, to make a significant pause between successive sentences, and to repeat the whole sentence if they made a mistake. The set of five sentences was read and

recorded at least twice for each participant, but the first reading was used for the analysis unless an uncorrected error occurred in one of the sentences. The experimenter did not provide any other indications about how to read the sentences. Recordings were made directly to computer using Praat (Boersma & Weenink, 2006) at a sampling rate of at least 22kHz.

2.4 Segmentation and analyses

The speech samples were segmented into vocalic and consonantal intervals by the first author. The segmentation was done by visual inspection of the waveform and spectrogram using Praat. Each vocalic interval comprised one or a sequence of consecutive vowels, irrespective of syllable, morpheme or word boundaries. Similarly, consonantal intervals were constituted of any number of consecutive non-vocalic segments.

As a general criterion, the start-point of a vocalic interval was taken as the onset of the first pitch period of the first vowel, and the end-point as the offset of the last pitch period of the last vocalic segment within the same interval. For more details about the specific criteria we used for segmentation, see White & Mattys (2007a).

Rhythm scores were derived for %V, VarcoV and rPVI_C, as described above. Scores were calculated for each of the five sentences produced by each of the six speakers in each language group: a total of 120 sentences (5 sentences x 6 speakers x 4 groups). The scores for the five target sentences were averaged for each speaker, and the score of each speaker used for by-subject analyses.

3. Results

Table 1 presents results for the three metrics for the two L1 and L2 groups, as well as the average speech rate for each group.

Table 1. Means (standard errors) of rhythm metrics for L1 and L2 speakers of Canadian English and Japanese

	L1 English	L2 English	L1 Japanese	L2 Japanese
	Eng _{ENG}	Eng _{JA}	Ja _{JA}	Ja _{ENG}
<i>Rhythm metrics</i>				
%V	47 (0.6)	46 (0.9)	54 (0.6)	54 (1.1)
VarcoV	52 (2.7)	46 (1.8)	56 (1.2)	54 (2.6)
rPVI_C	66 (1.4)	70 (3.3)	47 (3.5)	62 (4.2)
<i>Speech rate</i>				
Syllables/sec.	5.6 (0.2)	4.8 (0.2)	6.5 (0.2)	5.6 (0.2)

With regard to the relative scores of the L1s: as expected the %V score is higher for L1 Japanese speakers than for Canadian English speakers, reflecting the simplicity of Japanese syllable structure relative to Canadian English. This

factor is also reflected in the higher rPVI_C score for Canadian English, despite the single-geminate contrast in Japanese. As predicted, both Japanese and Canadian English demonstrate considerable variation in their vocalic intervals. This variation, as captured by VarcoV, appears equivalent for both languages, although in English the primary source of variation is the stressed-unstressed vowel contrast and in Japanese it is the long-short vowel distinction.

3.1 %V and VarcoV

Figure 1 shows the %V and VarcoV results for the L1 and L2 English and Japanese groups. The difference between the L1 groups is significant for %V ($t(10) = 7.936$, $p < .001$) but not for VarcoV ($t(10) = 1.647$, $p > .05$). The scores of the L2 speakers are comparable to those of native speakers for both metrics. That is, we did not find any significant difference on either %V or VarcoV between the L1 English speakers and Japanese speakers of L2 English, or between L1 Japanese speakers and English speakers of L2 Japanese.

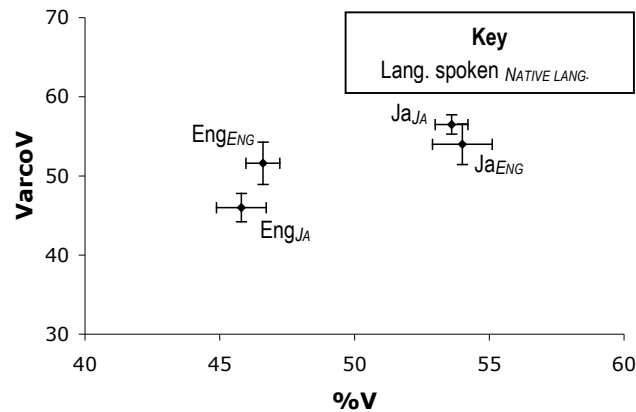


Figure 1. %V (x-axis), VarcoV (y-axis), and standard error bars for L1 and L2 speakers of Canadian English and Japanese.

To evaluate if Japanese speakers performed like native English speakers on the stressed-unstressed contrast, as suggested by the lack of difference in their VarcoV scores, we selected 11 pairs of consecutive stressed and unstressed syllables (listed in Table 2) from the sentences used for our experiment, and divided the duration of the stressed vowels by the duration of the unstressed vowels.

Table 2. List of syllable pairs used to calculate the ratio of the duration of stressed syllables to the duration of the unstressed syllables. Syllables within the same word are separated by a period.

Stressed-unstressed syllables	Unstressed-stressed syllables
<i>su.per</i>	<i>of poor</i>
<i>mar.ket</i>	<i>to make</i>
<i>mo.ney</i>	<i>the best</i>
<i>fa.mous</i>	<i>to pave</i>
<i>chair.man</i>	<i>co.mmi(ttee)</i>
<i>met this</i>	--

The stressed-unstressed ratios for L1 and L2 speakers of English are shown in Figure 2, indicating that, for native English speakers, the stressed vowels measured were 2.7 times longer than the preceding or following unstressed vowels. This ratio is 1.7 for native Japanese speakers speaking L2 English. Thus, these L2 speakers of English tend not to shorten unstressed vowels as much as L1 speakers. The difference between the L1 and L2 speakers of English is statistically significant ($t(10) = 4.157, p < .01$) and the effect size is also very large ($r = .8, p < .01$) indicating that even though Japanese speakers appear to perform like native English speakers according to VarcoV scores, they do not realize the stressed-unstressed contrast to the same degree as native English speakers. The variation in vowel duration captured by VarcoV may derive instead from Japanese speakers applying their native long vs short vowel duration contrast to analogous English vowels, such as tense-lax pairs.

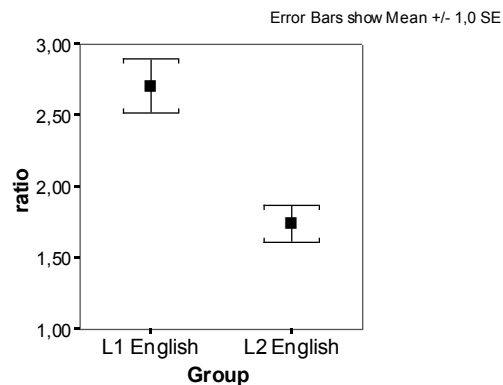


Figure 2. Ratio of the duration of stressed vowels divided by the duration of unstressed vowels for L1 and L2 speakers of English.

3.2 rPVI_C

Scores for rPVI_C, shown in Figure 3, are significantly higher for L1 English speakers compared with L1 Japanese speakers ($t(10) = 5.194$, $p < .001$). With regard to L2 groups, rPVI_C scores for Japanese speakers of L2 English (Eng_{JA}) are comparable to those of native English speakers ($t(10) = 1.203$, $p > .05$), whereas the scores of English speakers of L2 Japanese (Ja_{ENG}) are significantly higher than those of native Japanese speakers ($t(10) = 2.855$, $p < .05$).

As mentioned in the introduction, the rPVI_C metric can be affected by speech rate variation, and our L2 speakers indeed exhibited slower speech rates than our L1 groups, possibly affecting the scores obtained on this metric. In particular, the average speech rate for our native Japanese speakers (Ja_{JA}) is 6.5 syllables per second compared to 5.6 for our L2 Japanese speakers (Ja_{ENG}). This relatively small difference in speech rate may not fully explain the considerable difference obtained on the rPVI_C metric by the L1 and L2 Japanese groups.

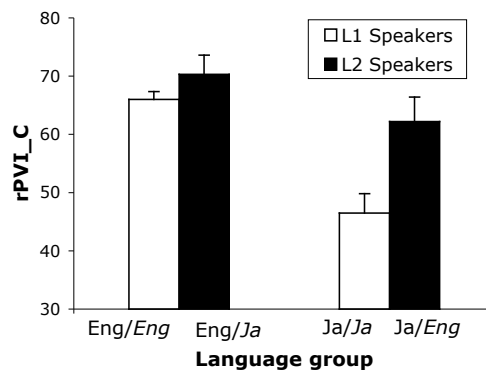


Figure 3. rPVI_C (y-axis), and standard error bars for L1 and L2 speakers of Canadian English and Japanese.

English speakers of Japanese, aware of the durational contrast between singleton and geminate consonants, but lacking it in their native language, might exaggerate the contrast in production by shortening single consonants, lengthening geminate consonants, or both. To test this hypothesis, we calculated a geminate-singleton ratio: the duration of geminate consonants in our corpus (five in total) divided by the duration of their singleton counterparts. The consonants used for this analysis are listed in Table 3.

Given that word length, word position and utterance length are not controlled between singleton and geminate samples, the resulting ratios do not reflect the true relative length of singleton versus geminate consonants in Japanese, which is also affected by differences in place and manner of

articulation. Rather, the comparison between L1 and L2 speakers indicates how Japanese-like the native English speakers are in their production of these consonants.

Table 3. List of singleton-geminate pairs used to calculate the geminate-singleton ratio. A period after a word indicates that the word appears in sentence-final position.

Sentence(s)	Single	Geminate
3 & 1	de <u>t</u> a.	hajim <u>tt</u> a.
2	e <u>n</u> o	Noom <u>in</u> <u>n</u> o
2	<u>t</u> amatta.	tam <u>tt</u> a.
4	<u>s</u> ooki	Shu <u>ss</u> ango
4	<u>t</u> akamatta.	takam <u>tt</u> a.

As shown in Figure 4, the singleton-geminate ratios indicate that English speakers did not exaggerate the contrast between singleton and geminate consonants in Japanese. The score of English speakers (1.9) is not significantly different from that of native Japanese speakers (2.1) ($t(10) = 1.1, p > .05$).

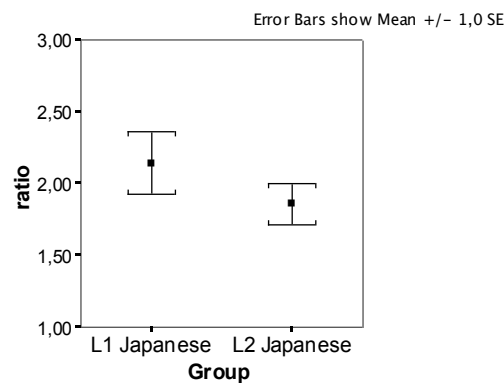


Figure 4. Geminate-singleton durational ratio for L1 and L2 speakers of Japanese.

The results for the singleton-geminate ratio indicate that an alternative explanation must be sought for the differences in rPVI_C scores for L1 and L2 Japanese speakers. Given that voiceless consonants are aspirated in English but not in Japanese, we hypothesized that English speakers may lengthen voiceless consonants in Japanese through aspiration. To investigate this possibility, we measured Voice Onset Time (VOT) of /t/ and /k/ in positions analogous to those in which they would be aspirated in English: i.e. word-initially or as the onset of a heavy (i.e. bimoraic or trimoraic) syllable. Table 4 indicates that Japanese speakers' mean VOT for these consonants was 35 ms, whereas among the native

English speakers, the mean VOT for the same consonants was 64 ms, almost twice as long. This difference is significant and the effect size is large ($t(12) = 3.608$, $p < .01$; $r = .7$, $p < .01$). The considerable difference in aspiration duration is likely to be partly responsible for the increased variation of the consonantal intervals among the native English speakers in L2 Japanese.

Table 4. Voice Onset Time (in milliseconds) of voiceless Japanese consonants by native Japanese speakers and native English speakers of L2 Japanese.

Japanese words	L1 Japanese (Ja _{JA})	L1 English (Ja _{ENG})
Containing /t/		
‡aiinshiteiku	24	82
‡akamatta	17	42
‡anoshimu	24	60
Containing /k/		
saik <u>en</u>	36	54
koo <u>z</u> ui	36	72
oo <u>k</u> ina	69	69
keiko <u>o</u> ga	37	71
Average	35	64

4. General discussion

Japanese speakers' rhythm production in L2 English appeared comparable to native Canadian English speakers on all metrics (%V, VarcoV and rPVI_C). Despite these results, however, we found that Japanese speakers did not realize, like native English speakers, the stressed-unstressed contrast that the VarcoV scores are intended to reflect.

English speakers' %V and VarcoV scores for L2 Japanese were comparable to those of native Japanese speakers, but they appeared to have greater variation in consonantal interval duration, with a score on rPVI_C closer to that of L1 English speakers. This difference appears to be unrelated to the possible L2 strategy of exaggerating the durational contrast between single-geminate consonants. However, L1 phonology seems to play a role in those results, as reflected in English speakers' aspiration of voiceless consonants in their Japanese production. Speech rate differences may also have affected rPVI_C scores.

The %V scores of our L2 speakers were comparable with those of the L1 groups, and did not reflect speakers' use of L1 phonotactics or phonological processes in their L2. For example, many Japanese speakers produced coda /n/ in English words as a nasalized vowel (e.g. in *chain*, *town*), while some used vowel epenthesis at the end of words (e.g. *standards* [standaru], *shopping* [ʃɔpiŋgu]) or elided some consonants in syllable-final position.

English speakers, on the other hand, did not apply common phonological processes in Japanese as much as native Japanese speakers. For instance, we found fewer occurrences of vowel devoicing/deletion than in our L1 native Japanese speakers samples, and coda /n/ was not as commonly produced as a nasalized vowel. It also remains to be seen whether English speakers produce the short-long vowel contrast in a way comparable to native Japanese speakers.

In sum, the metrics investigated (%V, VarcoV, rPVI_C) captured some correlates of speech rhythm and have undoubted application in comparisons of L1 and L2 production (see, for example, White & Mattys, 2007a, 2007b). However, the interpretation of rhythm scores is not straightforward, and should be considered as a guideline rather than evidence for native-like rhythmic proficiency, especially when applied to a quantity language, such as Japanese, which lacks a stress contrast.

Appendix A. Sentence materials

English

The supermarket chain shut down because of poor management.
Much more money must be donated to make this department succeed.
In this famous coffee shop they serve the best doughnuts in town.
The chairman decided to pave over the shopping centre garden.
The standards committee met this afternoon in an open meeting.

Japanese

Oono shigo ni machi no saiken ga hajimatta.
Noomin no sonchoo e no fuman ga tamatta.
Natsu no koozui de zuibun ookina higaiga deta.
Shussango sooki ni taiinshiteiku keikooga takamatta.
Konshuu mo uta bangumi o tanoshimu jikan ga nai.

Note: For the Japanese sentences, participants were given the choice between a version written in *Romaji*, *Hiragana*, or *Kanji*.

References

- Akamatsu, Tsutomu (1997). *Japanese phonetics: Theory and practice*, Newcastle: Lincom Europa.
- Boersma, Paul, & Weenink, David (2006). *Praat: Doing phonetics by computer* (version 4.3.04) [computer program]. Retrieved 21st March 2006, from <http://www.praat.org/>.
- Carter, Phillip M. (2005). Quantifying rhythmic differences between Spanish, English, and Hispanic English. In R. S. Gess, & E. J. Rubin (Eds.), *Theoretical and experimental approaches to romance linguistics: Selected papers from the 34th linguistic symposium on romance languages (Current issues in linguistic theory 272)* (pp. 63-75). Amsterdam, Philadelphia: John Benjamins.
- Dauer, Rebecca M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.

- Dellwo, Volker, & Wagner, Petra. (2003). Relations between language rhythm and speech rate. In *Proceedings of the 15th international congress of phonetics sciences* (pp. 471–474). Barcelona.
- Grabe, Esther, & Low, Ee Ling (2002). Durational variability in speech and the rhythm class hypothesis. In N. Warner, & C. Gussenhoven (Eds.), *Papers in laboratory phonology 7* (pp. 515-546). Berlin: Mouton de Gruyter.
- Gut, Ulrike (2003). Prosody in second language speech production: The role of the native language. *Fremdsprachen Lehren und Lernen*, 32, 133-152.
- Lin, Hua, & Wang, Qian (2005). Vowel quantity and consonant variance: A comparison between Chinese and English. Paper presented at the Conference *Between Stress and Tone*, Leiden, June 2005.
- Lloyd James, Arthur (1940). *Speech signals in telephony*. London: Pitman & Sons.
- Low, Ee Ling, Grabe, Esther, & Nolan, Francis (2000). Quantitative characterisations of speech rhythm: 'Syllable-timing' in Singapore English. *Language and Speech*, 43, 377-401.
- Nazzi, Thierry, Bertoncini, Josiane, & Mehler, Jacques (1998). Language discrimination by newborns: Towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 756-766.
- Ramus, Franck, Dupoux, Emmanuel, & Mehler, Jacques (2003). The psychological reality of rhythm classes: Perceptual studies. In *Proceedings of the 15th international congress of phonetic sciences* (pp. 337-342). Barcelona.
- Ramus, Franck, Nespor, Marina, & Mehler, Jacques (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.
- Tsujimura, Natsuko (1996). *An introduction to Japanese linguistics*, Malden: Blackwell Publishers.
- Vance, Timothy J. (1987). *An introduction to Japanese phonology*, Albany, State University of New York.
- White, Laurence, & Mattys, Sven L. (2007a). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35, 501-522.
- White, Laurence, & Mattys, Sven L. (2007b). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.), *Segmental and prosodic issues in romance phonology* (pp.237-257). Current issues in linguistic theory series. Amsterdam, Philadelphia: John Benjamins.
- White, Laurence, Mattys, Sven L., Series, Lucy, & Gage, Suzy (2007). Rhythm metrics predict rhythmic discrimination. In *Proceedings of the 16th international congress of phonetic sciences* (pp. 1009-1012). Saarbrücken.