말소리와 음성과학
Phonetics and Speech Sciences

Check for updates

Acoustic correlates of L2 English stress — Comparison of Japanese English and Korean English*

Takayuki Konishi · Jihyeon Yun · Mariko Kondo**

Abstract

This study compared the relative contributions of intensity, F0, duration and vowel spectra of L2 English lexical stress by Japanese and Korean learners of English. Recordings of Japanese, Korean and native English speakers reading eighteen 2 to 4 syllable words in a carrier sentence were analyzed using multiple regression to investigate the influence of each acoustic correlate in determining whether a vowel was stressed. The relative contribution of each correlate was calculated by converting the coefficients to percentages. The Japanese learner group showed phonological transfer of L1 phonology to L2 lexical prosody and relied mostly on F0 and duration in manifesting L2 English stress. This is consistent with the results of the previous studies. However, advanced Japanese speakers in the group showed less reliance on F0, and more use of intensity, which is another parameter used in native English stress accents. On the other hand, there was little influence of F0 on L2 English stress by the Korean learners, probably due to the transfer of the Korean intonation pattern to L2 English prosody. Hence, this study shows that L1 transfer happens at the prosodic level for Japanese learners of English and at the intonational level for Korean learners.

Keywords: L2 English prosody, lexical stress, Japanese learners, Korean learners

1. Introduction

Prosody plays an important role in speech communication. It also affects the perception of second language (L2) speech. While segmental accuracy is essential for a word to be recognized correctly, prosodic features carry a lot of information that helps the listener to understand parts of speech, syntactic structure, emphasized words, speech acts, as well as the speaker's emotions and attitudes. Therefore, using wrong prosody may significantly impair intelligibility and affect comprehensibility of L2 speech.

There are also empirical supports for the importance of L2 prosody, especially in comparison to L2 segments. Anderson-Hsieh *et al.* (1992) found that although both segmental and prosodic features influence significantly the judgement proficiency of L2 speech, prosody plays the more important role. Saito *et al.* (2016) reported that prosody affected the comprehensibility of the English speech by all levels of Japanese learners, whereas segmental accuracy only contributed to the comprehensibility of advanced learners' speech.

L2 phonology is usually affected by the phonology of the First Language (L1). English is a stress-accent language, with its stress being manifested by intensity (stressed syllables are louder than

unstressed syllables), F0 (stressed syllables are normally higher in pitch[1]) and duration (stressed syllables are longer); also, vowels in unstressed syllables are reduced to schwa (/ə/) sound, which can be observed by looking at the spectra (F1 and F2 values) of the vowels (e.g. Roach, 2009; Knight, 2012). Therefore, the accurate manifestation of L2 English stress depends on whether these four acoustic correlates are phonologically used in the learner's L1.

Japanese and Korean[2] have different phonological systems, which will have different influences on the L2 English of Japanese and Korean learners (hereafter JP learners and KR learners). Japanese has pitch-accent; both lexical accent and intonation are solely manifested by fundamental frequency (F0). The presence or absence of accent does not affect the quality of vowels, or durational reduction. This is because Japanese is a mora-timing language, and the duration of the mora is supposed to be more or less equal (e.g. Vance, 2008). In contrast, Korean is an intonation language. There is no lexical stress. Instead, the pitch of a syllable is affected by the intonation pattern at sentence level. Thus pitch accent is used to demarcate prosodic phrases in Korean (Jun, 2005; Lee et al., 2006). In addition, it is argued by some scholars that young Korean use pitch (i.e. F0) for the lexical distinction of aspirated and lenis stops (e.g. Silva, 2006; Kang, 2014). Another difference is that Japanese has lexical vowel length contrast while Korean does not use duration for any phonological contrast. Neither language has phonological contrast by intensity or vowel quality (<Table 1>).

Therefore, comparing L2 English by JP learners and KR learners will shed light on the mechanism of L1 phonological transfer to L2 prosody.

Table 1. Phonologically used parameters in each language

	Intensity	F0	Duration	Spectra
English	✓	✓	✓	✓
Japanese		✓	✓	
Korean		✓		

2. Previous studies

There are far fewer studies on L2 prosody than on L2 segments. The studies by Lee et al. (2006) and Konishi & Kondo (2015) are relevant to the current analysis.

Lee et al. (2006) investigated the lexical stress contrast by highly proficient KR learners and JP learners. They showed that KR learners had nativelike contrast of F0, but not intensity, duration or spectra. On the other hand, the JP learners had nativelike contrast of intensity, F0 and duration, but not spectra.

Konishi & Kondo (2015) revealed that, while highly proficient JP learners manifested nativelike durational contrast in the realization of L2 English stress, which is consistent with the results of Lee et al. (2006), less proficient JP learners had significantly different use of duration compared to that of the native English speakers (EN). However, even these less proficient JP learners had durational contrast between stressed and unstressed vowels.

This implies that there are certain stages in the acquisition of L2 English stress. In other words, even if there is a statistical difference between EN and learners in the use of an acoustic parameter, the learners may still be able to use the parameter contrastively. Therefore, the developmental change is the extent to which the learner uses the parameter to manifest L2 English stress.

By investigating the relative contribution of each constituent in manifesting L2 English stress, this study aims to investigate such developmental change, as well as the influence of the different L1s.

3. Data

The data for the analysis were extracted from the Asian English Speech cOrpus Project (AESOP) corpora, multinational L2 English speech corpora constructed in several Asian countries on the same platform (e.g. Visceglia et al., 2009). The data of the native EN group and JP learners (JP group) were extracted from the AESOP corpus in Japan (J-AESOP) and the KR learners' data (KR group) were extracted from the AESOP corpus in Korea (K-AESOP). The whole dataset consisted of the speech of 25 native speakers (EN group), 72 Japanese learners of English (JP group) and 139 Korean learners of English (KR group). All JP learners spoke Tokyo accent while the accents of the KR learners were diverse.

The analyzed speech was the read speech of eighteen 2-to-4-syllable words (money, apartment, misunderstand etc.; See <Table 2>) uttered in the carrier sentence *I say/said WORD five/ten times*.

Table 2. List of test words (The syllabification follows Longman Dictionary of Contemporary English, 6th edition. Syllables with primary stress are underlined.)

after**noon**, a**part**ment, a**vail**able, Cali**for**nia, de**part**ment store, **el**evator, ex**pe**rience, **hos**pital, infor**ma**tion, **Jan**uary, Japa**nese**, misunder**stand**, **mon**ey, **morn**ing, over**night**, **su**permarket, to**mor**row, **vid**eo, Vietna**mese**, white **wine**

The English proficiency of the JP group was rated using a separate recording of the learners reading *the North Wind and the Sun* from the Aesop's Fables. Their proficiency was rated by 2 native English and 6 non-native English phoneticians (<Table 3>) on a 10-point scale (in 1-point increments). The rating was done impressionistically with the criteria shown in <Table 4>. The inter-rater correlation coefficients ranged from 0.58 to 0.81. The proficiency of the KR group was not assessed because the K-AESOP corpus did not have a proficiency score.

---

[1] More precisely speaking, stressed syllables are loci of pitch accent (H* or L). In the current experimental design, in which the test words were imbedded in a carrier sentence, all the test words bear H* pitch accent, at least in the utterances of native English speakers.

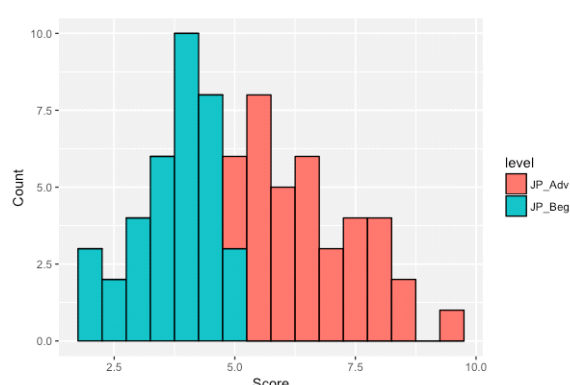[2] Tokyo Japanese and Seoul Korean, unless stated otherwise.

Table 3. Information about the JP group proficiency raters

	Native language	Degree	Field
1	English	MA	Phonetics
2	English	PhD	Speech science
3	Japanese	MA	Phonetics
4	Japanese	MA	Phonetics
5	Spanish	MA	Speech science
6	German	MA	Phonetics

Table 4. JP group proficiency rating criteria

| − Clear lexical stress |
| − Speech rhythm |
| − Speech free of wrong insertions/elisions of segment |

Using the median proficiency score (5.04), the JP group was divided into two proficiency groups: advanced group (JP_Adv, n =36) and beginner group (JP_Beg, n = 36). The distribution of the subjects is shown in <Figure 1>.

Figure 1. Proficiency score distribution of JP group learners
Red: JP_Adv / Blue: JP_Beg

4. Analysis

A logistic regression was conducted using *R 3.4.0*. The dependent variable was a Boolean value of whether the vowel was primary-stressed or not[3]. The independent variables were the four acoustic correlates of English stress, i.e. mean intensity (dB), mean F0 (in semitone; re = 1 Hz), duration (ms) and spectra of the vowel. Vowel spectra were converted to the Euclidean distance between each vowel either stressed or unstressed, and the center of the vowel space of each speaker. The vowel position was identified with F1 and F2 values, and the center of the vowel space was calculated by averaging both F1 and F2 values of all the canonical schwa sounds in the test words (e.g. the first and third <a> of <available>; <Table 5>). Rhotic shwas were excluded. To eliminate individual differences, each of the four parameters was z-score normalized.

Table 5. Schwa sounds of the test words

Test word	IPA
Apartment	/əpɑrtmənt/
Available	/əveɪləbəl/
California	/kæləfɔrnjə/
Department store	/dɪpɑrtmənt stɔr/
Elevator	/eləveɪtɚ/
Experience	/ɪskpɪriəns/
Hospital	/hɑspɪtəl/
Information	/ɪnfɚmeɪʃən/
Japanese	/dʒæpəniz/
Misunderstand	/mɪsəndɚstænd/
Tomorrow	/təmɑroʊ/

The other 9 words lack non-rhotic schwa (*afternoon, January, money, morning, overnight, supermarket, video, Vietnamese* and *white wine*). The phonetic transcriptions are based on the ones of J-AESOP corpus.

Firstly, a full-model was constructed with *glm()* incorporating all the main effects and the two-way interactions. Then, an optimal model was generated using the backward elimination method with the function *step()*.

For all groups, all the main effects were significant in determining whether the vowel was stressed. There were also significant interactions, with different combinations for each speaker group (<Table 6>).

Table 6. Significant interactions between acoustic correlates influencing vowel stress for each speaker group

Speaker group	Significant interactions
EN	− F0-spectra
JP_Adv	− intensity-F0
	− intensity-duration
	− duration-spectra
JP_Beg	− intensity-duration
	− F0-duration
	− duration-spectra
KR	− intensity-duration
	− intensity-spectra
	− F0-spectra
	− duration-spectra

Next, to investigate the relative contribution of each acoustic correlate, the coefficients of the main effects and interactions were assessed.

5. Results

The coefficients of intensity, F0, duration, spectra and interaction for the EN group were 0.85, 0.96, 1.53, 0.38 and 0.19, respectively (rounded to 2 decimal place; See <Table 7>). These values were then converted to percentages: 21.7, 24.6, 39.1, 9.6 and 4.9 (<Figure 2>). The relative contributions of each parameter to L2 English stress by each of the non-native groups, JP_Beg, JP_Adv and KR, are shown in <Figures 3-5>, respectively.

---

3 Vowels with secondary stress were excluded from the data analysis.

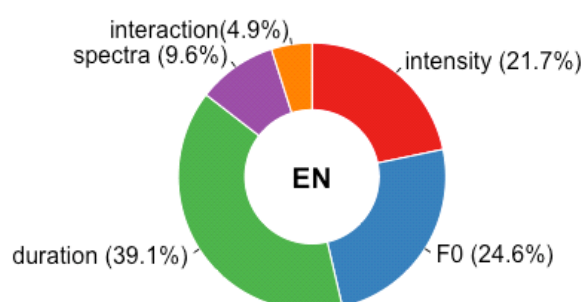**Table 7.** Coefficients of each acoustic correlate: EN

	Est	Std Err	*z* value	*p* value
(Intercept)	-0.73[1]	0.08	-8.68	$<2e^{-16}$
Int	0.85	0.10	8.75	$<2e^{-16}$
F0	0.96	0.11	8.94	$<2e^{-16}$
Dur	1.53	0.10	15.09	$<2e^{-16}$
Spec	0.38	0.08	4.66	$3.15e^{-16}$
F0:Spec	0.19	0.10	1.94	0.05

[1] All values rounded to 2 decimal place.

EN, English speakers; Est, Estimate; Std Err, standard error; Int, intensity; Dur, duration; Spec, spectra

A:B denotes interaction between A and B.

All four acoustic correlates contributed to the manifestation of stress in the EN group (<Figure 2>). Duration had the strongest influence (39.1%) on whether a vowel was stressed or not. In contrast, the influence of spectra (9.6%) was much lower, less than half the influence of intensity and F0 (21.7% and 24.6% respectively).
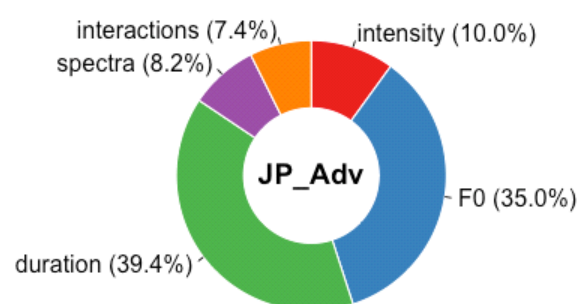

Figure 2. Relative contributions (%) of acoustic correlates of stress: EN group

The results of both JP groups (<Figure 3> and <Figure 4>) were quite similar except that there was more use of intensity by JP_Adv learners (10.0% in contrast to 6.0% by JP_Beg learners). What compensated for the difference in intensity was F0, which was used less by JP_Adv learners than by JP_Beg learners (35.0% vs. 41.8%). It is important to note that even the JP_Adv learners did not use as much intensity as the EN group (10.0% vs. 21.7%).
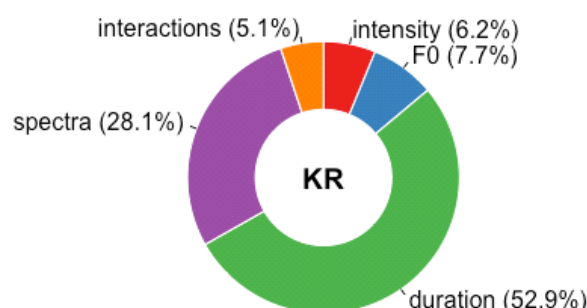

Figure 3. Relative contributions (%) of acoustic correlates of stress: JP_Beg group

Figure 4. Relative contributions (%) of acoustic correlates of stress: JP_Adv group
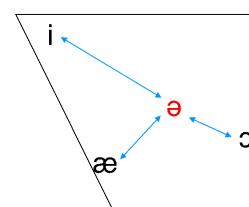
In the KR group (<Figure 5>), duration was the primary contributor to L2 English stress (52.9%), followed by spectra (28.1%). The influences of these two correlates were much larger than those of the EN group or either JP group. In contrast, both intensity and F0 had much less influence (6.2% and 7.7% respectively).

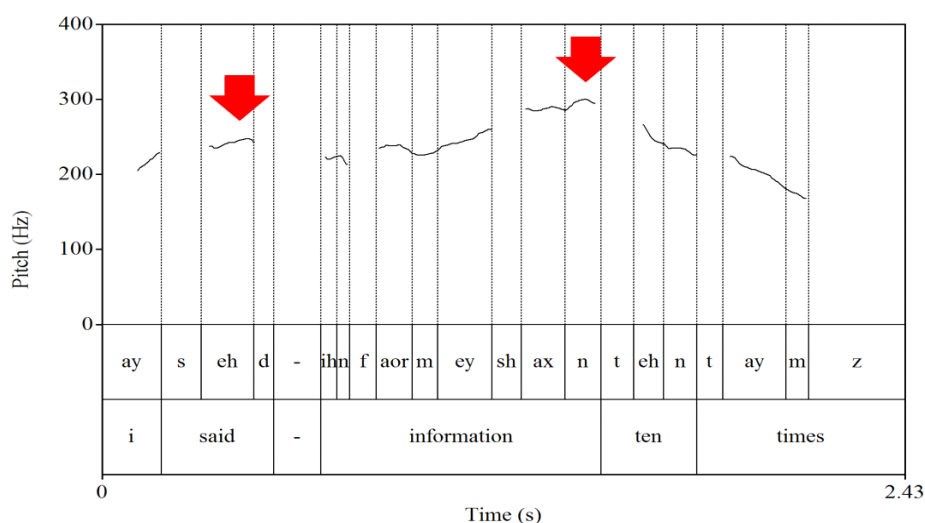**Figure 5.** Relative contributions (%) of acoustic correlates of stress: KR group

6. Discussion

Firstly, the weak influence of spectra on L1 English stress (9.6%) may have been due to the limitation of using Euclidian distance to measure the extent of reduction. F1 and F2 values do not depend solely on whether a vowel is reduced or not, but also on the relative position from the center of the vowel space (<Figure 6>). For example, /i/, /æ/ and /ɔ/ are all non-reduced vowels, but they all have different distances from the center of the vowel space, i.e. schwa. Hence, if they were judged solely by the Euclidian distance, /æ/ and /ɔ/ would be much more *schwa-like* than /i/.

Figure 6. Schematic representation of relative distance of each vowel from the center of the vowel space (i.e. schwa).
See /i/ is much farther from the schwa than are /æ/ and /ɔ/

Figure 7. Examples of phrase-final H tones in Korean English (indicated by the vertical red arrows)

On the other hand, the relatively strong influence of duration on L1 English stress might have been because of the diphthongs and rhotic vowels in the test words. In fact, the vowel in the stressed syllable of 50% (10/20) of the test words was either a diphthong or rhotic vowel.

Both JP groups showed the transfer of L1 phonology. As explained earlier, Japanese uses F0 and duration to manifest phonological contrasts. This probably explains the large effects of F0 and duration in manifesting L2 English stress (35.0% and 39.4% respectively for JP_Adv group; and 41.8% and 36.5% for JP_Beg group).

The main differences between the JP_Beg learners and JP_Adv learners were that the advanced learners used more intensity (10.0%) than the beginner learners (6.0%), and less F0 (35.0% vs. 41.8%). Hence, the developmental change as learners advance from beginner to advanced seems to be less reliance on F0, which is typically used to manifest Japanese pitch accent and intonation, and more use of intensity, which is not used in Japanese phonology but is used by EN to manifest stress.

The results of the KR group were different from those of Lee *et al*. (2006), who reported that F0 was the only parameter used by Korean learners to manifest lexical stress in English. In the current study, the large influence of duration and spectra by the KR group cannot be explained in terms of phonological transfer since they are not used in the learners' L1.

The weak effect of F0 (7.7%) found in our study could be due to the transfer of L1 Korean prosody. Accentual phrase in Korean is realized much more frequently with phrase final H tone than with L tone (Jun, 2000). In our study, the final syllable of the test words was often realized with phrase-final H tone (<Figure 7>). In other words, in our study it seems that the L1 intonational pattern was transferred for the manifestation of L2 prosody, rather than either L1 segmental phonology or lexical prosody. The same tendency was reported by Kang *et al*. (2012).

In addition, similar to the EN group, the impact of duration by the KR group in our study may be due to different syllable weights.

Lastly, it is important to note that, in contrast to the KR group in the study by Lee *et al*. (2006), all of whom were "Korean-English bilinguals" with at least 10 years of residence in US, the KR learners

in K-AESOP had varying proficiency, ranging from beginner to advanced levels. Even though there were no notable differences between JP_Adv group and JP_Beg group, we cannot assume that there would equally be no differences between advanced learners and beginner learners in the KR group. Therefore, rating proficiency levels in the K-AESOP corpus may help to explain why the current results were not consistent with those of the previous study.

7. Conclusions

The current study has demonstrated that the different realizations of L2 English lexical stress by JP learners and KR learners are based on different kinds of L1 transfer. For JP learners, the transfer was L1 lexical prosody, whereas for KR learners it was L1 intonation pattern.

For the JP group, the developmental change from beginner to advanced seems to be increased use of intensity, which is not important for L1 Japanese accent manifestation, and less reliance on F0 which is used in L1 Japanese phonology.

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, 529-555.

Jun, S. (2000). K-ToBI (Korean ToBI) labelling conventions. Version, 3.1. *In Proceedings of the KSPS Conference, The Korean*

Society of Phonetic Sciences and Speech Technology.

Jun, S. (2005). Korean intonational phonology and prosodic transcription. *Prosodic Typology: The phonology of intonation and phrasing* (pp. 201-229). Oxford: Oxford University Press.

Kang, S., Kang, J., & Kim, K. (2012). The phonetic realization of English unstressed vowels produced by Korean advanced learners: A comparative study of English words and English loanwords. *Phonetics and Speech Sciences*, 4, 3-11.

Kang, Y. (2014). Voice onset time merger and development of tonal contrast in Seoul Korean stops: A corpus study. *Journal of Phonetics*, 45, 76-90.

Knight, R-A. (2012). *Phonetics: A coursebook.* New York: Cambridge University Press.

Konishi, T., & Kondo, M. (2015). Developmental change in English stress manifestation by Japanese speakers. *In Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS XVIII).*

Lee, B., Guion, S., & Harada, T. (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese Bilinguals. *Studies in Second Language Acquisition*, 28, 487-513.

Longman Dictionary of Contemporary English (6th edition). (2014). Harlow: Pearson Education.

Roach, P. (2009). *English phonetics and phonology: A practical course* (4th edition). Cambridge: Cambridge University Press.

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37, 217-240.

Silva, D. (2006). Acoustic evidence for the emergence of tonal contrast in contemporary Korean. *Phonology*, 23, 287-308.

Vance, T. (2008). *The sounds of Japanese with audio CD.* New York: Cambridge University Press.

Visceglia, T., Tseng, C., Kondo, M., Meng, H., & Sagisaka, Y. (2009). Phonetic aspects of content design in AESOP (Asian English Speech cOrpus Project). *Proceedings of the 2009 Oriental COCOSDA International Conference on Speech Database and Assessments* (pp. 60-65). Urumqi.

● **Takayuki Konishi**
Graduate School of International Culture and
Communication Studies
Waseda University
1-6-1 Nishi-Waseda, Shinjuku-ku
Tokyo 169-8050, Japan
Tel: (+81)-3-5286-1728
Email: tkonishi@aoni.waseda.jp
Fields of interest: Phonetics, Second Language Acquisition

● **Jihyeon Yun**
Graduate School of Science and Technology
Sophia University
7-1 Kioi-cho, Chiyoda-ku
Tokyo 102-8554, Japan
Tel: (+81)-3-3238-3417
Email: yunjihyeon@gmail.com
Fields of interest: Phonetics

● **Mariko Kondo** corresponding author
School of International Liberal Studies &
Graduate School of International Culture and
Communication Studies, Waseda University
1-6-1 Nishi-Waseda, Shinjuku-ku
Tokyo 169-8050, Japan
Tel: (+81)-3-5286-1383
E-mail:mkondo@waseda.jp
Fields of interest: Phonetics, Language Acquisition,
Speech Science