English Lexical Stress in Awareness and Production: Native and Non-native Speakers

Mariko Sugahara[1], Sylvain Coulange[1, 2], and Tsuneo Kato[1]

[1]Doshisha University (Japan), [2]University Grenoble Alpes (France)

This study aims to clarify how the consistency between English primary stress patterns of which speakers are aware and those inferred from the acoustic properties in their speech differs among English native speakers (ENS), Tokyo/Keihan Japanese learners of English (JLE) and Seoul Korean learners of English (KLE), and considers the effect of their L1. English is a stress accent language in which word-initial primary stress is predominant ([1]). It also has subsidiary stress, too. Japanese is a pitch accent language in which about 14% of homophones are distinguished by pitch accent ([2]) and no subsidiary accent exists. The preferred accent position of morphologically complex and longer words in Japanese is not word-initial but rather word-medial ([3]). Seoul Korean, on the other hand, lacks lexically distinctive accent ([4]). We clarify how these differences in lexical accent systems among the three native languages affect the point mentioned above.

We conducted a paper-based stress-assignment task and a production task with 12 ENS, 14 JLE and 11 KLE. Although the paper-based task is a rehash of [5], the production task is new to this study. 19 triplets of morphologically related words (57 words) were adopted: three-syllable verbs ending with *ate*/*ute* whose prescriptive main stress falls on σ1 (e.g., *DO.mi.nate*), inflected verbs with the "stress-neutral" suffix *-ing* (e.g., *DO.mi.na.ting*), and derived nouns with the "stress-shifting" suffix *-ion* where σ3 is stressed (e.g., *do.mi.NA.tion*). In the paper-based task, the participants were asked to place a stress mark above the vowels of syllables that they considered primarily stressed. In the production task, they read aloud the 57 words embedded in carrier sentences, which was followed by forced segmental alignment (MFA3.0 by [6]) and acoustic analyses utilizing PLSPP (the Pauses and Lexical Stress Processing Pipeline by [7]). For each vowel interval, PLSPP first took (a) the mean of F0, (b) the max intensity value and (c) the duration. It secondly obtained the average of the speaker-normalized percentile values of (a), (b) and (c) for each vowel interval. The average score is called GS (Global Score) henceforth. Finally, it estimated the syllable containing the vowel with the highest GS to be acoustically most prominent.

The results of the paper-based stress-assignment task are summarized in Fig. 1. The stress assignment by ENS was predominantly prescriptive for all the three forms. JLE preferred the prescriptive stress patterns for the simple and the *-ion* forms, whereas they were divided into σ1-stress and σ3-stress for *-ing*, which may be because they are influenced by their L1 preferring word-medial accent in morphologically complex words. KLE's pattern was more varied with "zero" stress and σ2-stress, which implies that KLE have difficulty in acquiring English lexical stress patterns: they have set up the lexical accent parameter so that lexical accent is not encoded in the phonological representation ([8]) because their L1 lacks lexically distinctive accent.

Fig. 2 summarizes PLSPP's prominence estimation. Fig. 3 shows the ratio of agreement between PLSPP's estimation and the stress-assignment in the paper task. As for ENS, despite their uniformly prescriptive performance in the paper task for all the three forms, only the simple and the *-ing* forms with σ1-stress obtained PLSPP's prominence estimation that matched their stress-assignment: about half of the *-ion* forms with σ3-stress were estimated to have σ1-prominence. ENS, however, still distinguished *-ion* from *-ing* by the GS of σ3 even when they were equally estimated to have σ1-prominence by PLSPP (Fig. 4). It is because σ1 carries secondary stress in the *-ion* forms, and the secondarily stressed σ1 may have become acoustically more prominent than the mainly stressed σ3 due to analogy from the fact that word-initial main stress is predominant in English ([1]). In contrast, as for JLE's simple and *-ion* forms, for which they assigned prescriptive stress in the paper task, PLSPP's prominence estimation was in agreement with the prescriptive stress patterns regardless of whether they are of σ1-stress or σ3-stress (Figs. 2-3). That is, unlike ENS, they uniformly gave greater acoustical highlight to σ3 than to σ1 in the *-ion* forms. It is not surprising, considering that JLE is used to highlight only one syllable in their L1. KLE's ratio of agreement between the acoustically prominent and the stress-assigned syllables was low across the three forms (Fig. 3), which is consistent with the discussion above.

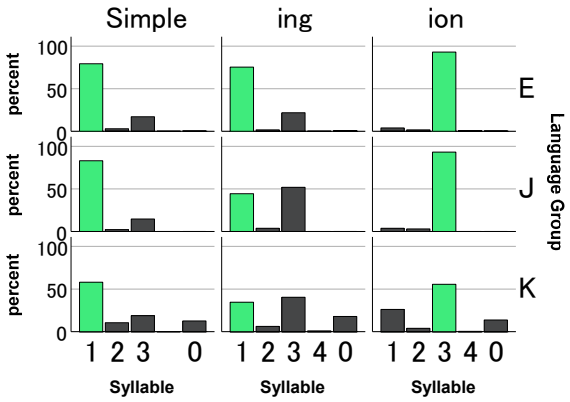

Fig. 1. Stress-assigned syllables (results from the paper-based stress-assignment task)
Note: "Simple" means the verb forms without suffixes. The numbers on the *x* axis represent syllables (e.g., "1" means that the stress mark was assigned to σ1, and likewise for the rest of the syllables). "0" means no stress was assigned to any of the syllables. The syllables with green bars are of prescriptive stress.

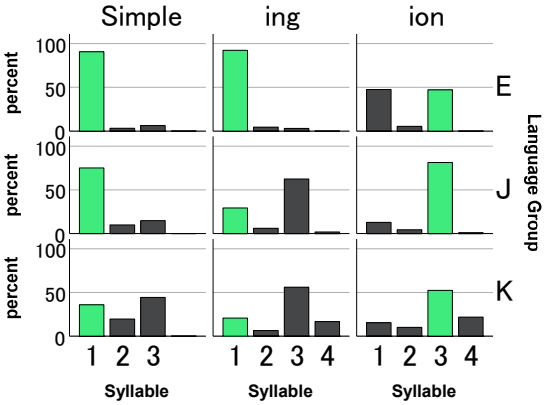

Fig. 2. Syllables estimated to be acoustically most prominent by PLSPP (results from the production task)
Note: The numbers on the *x* axis represent syllables (e.g., "1" means that σ1 was estimated to be acoustically most prominent, and likewise for the rest of the syllables). The syllables with green bars are of prescriptive stress.

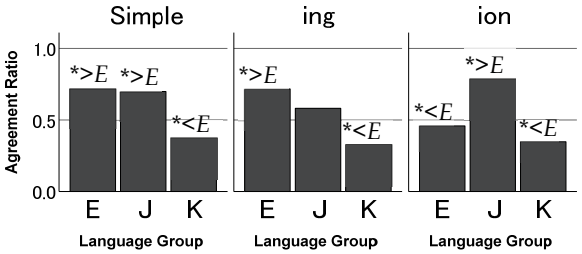

Fig. 3. The ratio of agreement between the stress-assignment (paper-based task) and PLSPP's prominence estimation (production task)
Note: "*<E*" and "*>E*" mean significantly smaller and significantly greater than the expected respectively according to the residual analyses following *chi*-square tests ($\chi^2(2)=66.73$, $p<.001$ for *Simple*; $\chi^2(2)=66.21$, $p<.001$ for *ing*; $\chi^2(2)=101.6$, $p<.001$ for *ion*).

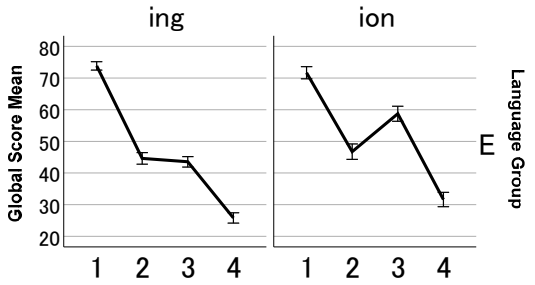

Fig. 4. The mean GS of the syllables in the *-ing* and the *-ion* forms when they were estimated to have σ1-prominence by PLSPP (ENS only)
Note: The numbers on the *x* axis represent syllables, and error bars show 95% confidence intervals.

References

[1] Cutler, A., & D. M. Carter. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language,* 2: 133-142.

[2] Sibata, T. & R. Sibata. (1990). To what extent can accents distinguish homophones? *Keiryoo Kokugogaku* 17(7): 311–323.

[3] Kubozono, H. (2006). Where does loanword prosody come from? A case study of Japanese loanword accent. *Lingua,* 116, 1140-1170.

[4] Jun, S-A. (1996). *The phonetics and phonology of Korean prosody*. Garland Publishing.

[5] Sugahara, M. (2019). The effect of native languages on the judgement of primary stress locations in English: a questionnaire study on stress assignment to present participle/gerundive forms with *-ing* and derived nominals with *-ion*. *Doshisha Studies in English*, 100: 165-221.

[6] McAuliffe, M., M. Socolof, S. Mihuc, M. Wagner & Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. In *Proceedings of the 18th Conference of the International Speech Communication Association*.

[7] Coulange, S. (n.d.). Pauses & lexical stress processing pipeline. https://ttk.gricad-gitlab.univ-grenoble-alpes.fr/lidilem/plspp

[8] Peperkamp, S. & E. Dupoux. (2002). A typological study of stress 'deafness'. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp.203-240). Mouton de Gruyter.