ISSN 0286-1291
Dōshisha Daigaku
Eigo Eibungaku Kenkyū

同志社大学
英語英文学研究

97号

2016年11月

同志社大学人文学会

同志社大学英語英文学研究 97号

同志社大学人文学会

DOSHISHA STUDIES IN ENGLISH

No. 97

November 2016

THE LITERARY ASSOCIATION
DOSHISHA UNIVERSITY

Is the Perception of English Stress by Japanese Listeners Influenced by the Distribution of Accent in their L1? In the Case of Truncated Word Stimuli

Mariko Sugahara

Abstract

The main goal of this study is to investigate whether Japanese listeners' knowledge of pitch accent distribution in their native language affects their auditory perception of English stress by comparing the response patterns of Japanese listeners with those of native English listeners and Seoul Korean listeners in a forced choice identification experiment using English truncated word stimuli. The location of accent in Japanese is skewed to the syllable that contains the antepenultimate mora, and the hypothesis tested in this study is that Japanese listeners are biased towards antepenultimate-mora stress even when perceiving English stress. The experiment consisted of two tasks. One of them used truncated stimuli consisting of only initial syllables extracted from noun-verb pairs with different stress patterns such as *TRANS*- from *TRANSplant* and *trans*- from *transPLANT* (Group 1). The results of the task did not support the hypothesis because all the three language groups showed the same response patterns. The other task used truncated stimuli extracted from four-syllable suffixed words, e.g., *DOminating* and *domiNAtion* (Group 2). The results obtained from the task with three-syllable truncated stimuli, e.g., *DOmina*- and *domiNA*-, in a condition where the two stress patterns were not distinguished by pitch complied with the prediction made by the hypothesis: the Japanese listeners showed more preference for the -*ion* forms, e.g., *domiNAtion*, that have stress on the antepenultimate mora than the other two language groups. In the final conclusion, the possibility that what matters in Japanese listeners' perception of English stress is in fact not the antepenultimate accent rule but the Latin accent rule is also considered.

1. Introduction

The main goal of this study is to investigate whether Japanese listeners' knowledge of lexical accent distribution in their native language affects their perception of English stress by carrying out a forced choice identification experiment using English truncated word stimuli. The location of accent in Japanese, e.g., Tokyo Japanese and Osaka-Kyoto Japanese, is skewed to the syllable that contains the antepenultimate mora (McCawley, 1968; Kubozono, 2006; among others), and this research aims to examine if Japanese listeners are affected by the antepenultimate accent even when perceiving English stress.

There are a substantial number of research results on suprasegmental properties in native languages or first languages (L1) affecting the perception of suprasegmental features in foreign languages or second languages (L2). Listeners are known to be sensitive to suprasegmental features such as pitch in foreign languages or L2 that are also used in their L1 as a lexical tone, lexical accent or lexical stress. For example, Beckman (1986) reports that Japanese listeners heavily rely on pitch to detect English stress while native English listeners evenly use not only pitch but also duration, amplitude and spectral quality for stress jusdgment. It is because the Japanese language utilizes only pitch to mark lexical accent and therefore Japanese listeners are sensitive to pitch information when detecting lexical stress in English. Ou (2010) considered Taiwan Mandarin listeners' perception of English stress. She found that they were able to correctly identify English stress when it was signaled by high pitch followed by a falling contour while they were not good at doing so when stressed syllables had a low rising contour. Her interpretation of the result is that Taiwan Mandarin listeners interpret stress in terms of tonal features. It is because they are influenced by the tonal system of their L1: Mandarin is a lexical tone language in which words are distinguished by tonal shapes. Ou & Guo (2015) reports that not only Taiwan Mandarin listeners but also Japanese listeners took high-falling pitch but not low-rising pitch as a cue for English lexical stress in a lexical decision task. This may be because Japanese accented syllables are always signaled by a high pitch immediately followed by a falling contour. Schaefer and Darcy (2014) compared the perception of Thai tones among the listeners of Mandarin (lexical tone language), Japanese (lexical pitch accent language), English (lexical stress language), Kyungsang Korean (lexical pitch accent language), and Seoul Korean (with neither lexical stress nor lexical accent). They report that the order of accuracy rate among the listeners of those language groups were as follows: Mandarin > Japanese = Kyungsang Korean > English > Seoul Korean. Their interpretation of the result is that the 'pitch functionality' of L1, i.e., 'the degree to which pitch differentiates lexical items in the L1,' affects the perception of foreign tones (Schaefer and Darcy, 2014: p.512). According to them, the pitch functionality of lexical tone languages is maximal because the specification of different shapes of pitch is made at the syllable level, that of pitch accent languages is intermediate because pitch specification is made at a wider domain, i.e., a word, and that of lexical stress languages is also intermediate but lower than that of pitch accent languages because pitch is not the only correlate of stress, and that of languages with neither lexical pitch accent nor stress is the lowest.

There are also studies on the influence of native lexical accent/stress distribution on the perception of foreign lexical accent/stress. In their production and perception study, Guion, Harada & Clark (2004) compared stress assignment to English disyllabic pseudo words among native English speakers, early Spanish-English bilinguals who had started to be massively exposed to English in the United States by the age of 6 and late Spanish-English bilinguals who had started to be massively exposed to English in the United States after the age of 15. They found that late Spanish-English bilinguals were more strongly biased towards initial-syllable stress than the other two groups of speakers. One of the possible interpretations of the result is that late bilinguals transferred the stress pattern in their L1. In Spanish, there are two regular stress patterns for disyllabic words: initial or final. Initial stress appears when the final syllable ends with a vowel, and final stress appears when the final syllable ends with a consonant. Among them, the initial stress pattern is more frequent because most of the disyllabic words in Spanish ends with a vowel. This distributional knowledge about their L1 stress was used even when the late bilinguals assigned stress to English disyllabic pseudo words. It is also reported by Guion (2005) in her production and perception study using the same English disyllabic pseudo words as those used in Guion et al. (2004) that late Seoul Korean-English bilinguals did not distinguish the stress pattern of nouns and that of verbs as much as native English speakers and early Seoul Korean-English bilinguals did. Guion states that it is because their L1 lacks lexical stress and they pay less attention to stress pattern differences between the two categories when learning English vocabulary.

Sugahara (2016a) posed the question whether the most frequent Japanese accent pattern, i.e., the antepenultimate accent, has any influence on their judgement of English stress locations in an auditory identification task. In that study, the response patterns of three groups of listeners were compared: native English listeners, Japanese learners of English mostly from the

Kansai area (Japanese listeners, henceforth) and Seoul Korean learners of English (Seoul Korean listeners, henceforth). In that perception task, the participants were asked to detect the syllable with primary stress in English disyllabic nouns and verbs forming minimal pairs in terms of the location of primary stress, such as *TRANSplant* (noun, trochaic) and *transPLANT* (verb, iambic). The final syllable of those pairs of words consisted of three morae: CVCC or CVVC. In one of Sugahara's (2016a) experimental conditions, the pitch of stimuli was synthesized so that those two types of words were not disambiguated by pitch. The main results obtained from that experimental condition are (a) the native English listeners showed no bias towards either initial or final stress, (b) both the Japanese and the Seoul Korean listeners preferred final stress, and (c) the Japanese listeners' final stress bias was significantly stronger than the Seoul Korean listeners'. Sugahara (2016a) interpreted the results as follows. First, the native English listeners were able to use acoustic information other than pitch to disambiguate the two stress patterns. Secondly, there was something in the stimuli which led both the Japanese and the Seoul Korean listeners to prefer more final stress than initial stress, e.g., the durational difference between the initial and the final syllables (see Section 3.2 of this article for more discussion). Nonetheless, the Japanese listeners showed stronger final-stress bias than the Seoul Korean listeners because the final syllable vowel in those stimuli was an antepenultimate mora and the Japanese listeners considered it as the most ideal location for stress being influenced by the predominance of antepenultimate-mora accent in their L1 lexicon.

The current study aims to support the conclusion made by Sugahara (2016a) by carrying out auditory perception tasks using truncated word stimuli. The tasks reported in this article and the task reported in Sugahara

(2016a) are from a single large experiment consisting of multiple tasks, and therefore participants and experimental environments are shared by the current study and Sugahara (2016a).

The organization of this article is the following. In Section 2, an overview of the lexical prosodic systems of English, Japanese and Seoul Korean is provided. It will be also explained here how their prosodic systems make different predictions for the perception patterns of English stress. Section 3 presents the result of the task with truncated stimuli consisting of only initial syllables extracted from noun-verb pairs such as *TRANS*- from *TRANSplant* and *trans*- from *transPLANT*. This stimuli set is called Group 1. The results obtained in the task did not support the hypothesis. Section 4 presents the result of another set of tasks with truncated stimuli, Group 2. The Group 2 stimuli set consisted of two-syllable and three-syllable stimuli extracted from four-syllable suffixed words with different stress patterns, e.g., *DOmi*- and *DOmina*- from *DOminating* and *domi*- and *domiNA*- from *domiNAtion*. The results obtained from the task with three-syllable stimuli complied with the prediction made by the hypothesis. Finally, concluding remarks will be provided in Section 5. In that section, the possibility that what matters in Japanese listeners' perception of English stress is in fact not the antepenultimate accent rule but the Latin accent rule is also considered.

2. Lexical stress and accent systems in the three languages

A brief overview of the lexical stress or accent systems in the three languages and some consideration to how those different systems may result in different sensitivity to lexical prominence are given in this section. Although some content of this section may overlap with the summary of the three languages' prosodic systems presented in Sugahara (2016a), it is necessary to have such overlap here to reexamine and reinterpret some of the ideas presented in the 2016a study.

Japanese and English are different in that the former is a pitch accent language while the latter is a stress language. They are, however, similar in a way that prosodic prominence (accent or stress) is distinctively used in the phonemic representation of words. Furthermore, the stress/accent distribution of those languages is skewed to certain positions in a word: initial-syllable (antepenultimate or penultimate-‘syllable’) stress is the most common in English and antepenultimate-‘mora’ accent in Japanese. Seoul Korean, on the other hand, has neither lexical stress nor lexical accent, and all seeming stress-related phenomena are at a phrase-level. The hypothesis tested in this study is that those typological differences regarding the distribution of stress/accent among the languages result in different patterns of English stress perception.

2.1. English

2.1.1. The distribution of English stress

English is a stress language in which the lexical prominence of a word is acoustically expressed with not only pitch[1] but also duration, vowel quality and amplitude (Beckman 1986, among others). At the same time, it is a ‘free stress’ language where stress locations are determined word by word, and every content word has a syllable with primary stress. Therefore, there exist minimal pairs whose meanings are distinguished by only stress locations, e.g., *díffer* vs. *defér*, *fórearm* (noun) vs. *foreárm* (verb), etc.[2] The distribution of stress is not completely random, however.

First, English primary stress positions are usually confined to one of the last three syllables of a word. Because monosyllabic, disyllabic and trisyllabic words are the majority in the English vocabulary, this tendency leads to the predominance of the initial stress in the English lexicon (Cutler & Carter, 1987: Clopper, 2002). Grammatical categories also matter when determining which of the last three syllables bears stress: verbs gravitate to final stress more than nouns. Based on the numerical figure given by Hammond (1999: 194), the following becomes clear. Among disyllabic words, 83% of nouns have primary stress on the penultimate (initial) syllable while 52% of verbs have primary stress there. When it comes to trisyllabic words, 50% of nouns have antepenultimate (initial) main stress whereas the proportion of antepenultimate main stress in verbs is only 20%.[3] Although the proportion of nouns and that of verbs having penultimate (middle) primary stress are the same (about 40%)[4], final stress is more dominant in verbs (40%) than in nouns (9%).[5] Since the number of verbs is only about half the number of nouns in English (Hammond, 1999:194), final-syllable stress is rarer than antepenultimate and penultimate stress, which also leads to the predominance of initial stress. Cutler & Carter (1987) reports that about 60 % of English vocabulary and about 90 % of all lexical tokens in spontaneous speech are those with word-initial primary stress.

Secondly, derivational suffixes also affect stress patterns. Although all inflectional suffixes do not change the original stress pattern of the stem to which they attach, derivational suffixes are divided into two groups depending on whether their attachment to a stem keeps or alters the original stress pattern of the stem. Examples of stress-neutral suffixes are -*ness* and -*less*, and those of suffixes that call for deviation from the original stress pattern of the stem are -*al*, -*ion*, -*ese*, -*eer*, -*ate*, -*oid*. The former are called ‘Class II’ suffixes and the latter ‘Class I’ suffixes.[6] For example, -*ion* requires the final syllable that contains the suffix to be extrametrical and a weight sensitive main stress foot to be created immediately before it: *va*.(*cá*)<*tion*>. When it comes to the suffix -*ate*, it receives secondary stress and a weight-insensitive main stress foot is created immediately before it: (*dé.sig*)(*nàte*). One of the experimental tasks of this study uses words that contain both -*ate* and -*ion*, e.g., *domination*. The base form of *domination* is the verb *dominate*, which also contains a bound morpheme -*ate*. When the verb base is bare without any derivational suffix, it has initial stress: (*dó.mi*)(*nàte*). Once the suffix -*ion* is attached, the main stress moves from the initial to the penultimate syllable: (*dò.mi*)(*ná*)<*tion*>. The task also uses verbs with an inflectional suffix -*ing*, e.g. *dóminàting*, which maintains initial stress.

Another thing to note is that prefixes such as *in*-, *re*-, *mis*-, *trans*-, are more likely to lack main stress according to the author's count of words stored in the CELEX database (Baayen, Piepenbrock & Gulikers, 1995). For example, the number of word lemmas which start with *TRANS*- (main stress) is only three and their total frequency is 1,460. On the other hand, the number of those which start with *trans*- (without main stress) is thirty-four and their total frequency is 2,676.

Those characteristics of English stress will be considered later again when making predictions for native English listeners' responses in the experiment carried out for this study.

2.1.2. Stress perception by native English listeners and initial stress bias

When native English listeners are not fully sure about the identity of English words they have heard, they are likely to respond that what they have heard is of initial stress. That is, they tend to rely on the knowledge of stress distribution when segmental and supersegmental content of English word stimuli is not clear enough.

Van Leyden and van Heuven (1996) made native English listeners hear both disyllabic or trisyllabic English words and the initial fragments of those words (e.g., only the initial syllable), and asked them to write down what they thought to be the ones they had heard. Interestingly, the majority of their error responses had initial stress regardless of the stress patterns of the original stimuli. Van Leyden and van Heuven argue that it is because the native English listeners gravitated to the most common stress pattern in English. Errors occur when participants are not sure about the identity of the stimuli, and in that kind of environment they might rely on their knowledge about English stress distribution.

Cooper, Cutler and Wales (2002) made native English listeners hear truncated stimuli consisting of only initial syllables extracted from words such as *MU*- from *MUsic* (initial stress) and *mu*- from *muSEum* (non-initial stress) in one of their experiments. They asked the listeners to identify which word they had heard, *MUsic* or *muSEum*. Their correct response rate was higher for the initial-stress words (more than 70%) than for the non-initial-stress words (less than 50%). That is, they were more biased towards initial stress. Cooper *et al*. also claim that the response patterns of their native English listeners were influenced by the predominance of initial stress in English. Since the stimuli consisted of only initial syllables in their task, both the segmental and the suprasegmental content of the stimuli were scarce. Therefore the listeners had to rely on their distributional knowledge of English stress when giving their responses.

The results obtained in Sugahara (2011) and those obtained in Sugahara (2016a) also imply that native English listeners' bias towards initial stress emerges when stimuli provided to them lack suprasegmental information to distinguish stress patterns. Sugahara's (2011) forced choice auditory experiment conducted previous to her 2016a study used trochaic noun-iambic verb minimal pair stimuli produced by a female native speaker of American English, such as *TRANSplant* and *transPLANT*. In the experiment, native English listeners gave more correct responses to initial-stress stimuli than to final-stress stimuli when the pitch track of the stimuli was made flat so that it did not disambiguate the two stress patterns. Sugahara (2016a) tried to replicate the result in the 2011 study, with similar minimal pair stimuli produced by a male native speaker of American English. The 2016a study, however, did not obtain initial-stress bias in native English listeners' responses in spite of the fact that pitch was synthesized in the same way as in Sugahara (2011). The difference may depend on the nature of stimuli. In fact, more durational contrast between the initial-stress stimuli (*TRANSplant*) and the final-stress stimuli (*transPLANT*) was available in the 2016a study than in the 2011 study. It is because when the stimuli were produced in the 2011 study, the words were deeply embedded in long sentences without any diacritic on the syllable with main stress when they were recorded. Since the sentential contexts were abundant to discriminate between the initial-stress nouns and the final-stress verbs, the speaker did not have to spend much effort to acoustically distinguish the two forms. Therefore, the two types of stimuli had similar durational properties. Once those stimuli were synthesized so that their noun forms and the verb forms are not disambiguated by pitch, they sounded so similar that the native English listeners had to guess the stress pattern of a given stimulus relying on their knowledge about the stress distribution in English. In the 2016a study, however, the words were visually presented to the native speaker embedded in a short sentence frame that did not provide any clue about their grammatical category and their stress pattern. Furthermore, in order to let the speaker know the stress pattern of the words, the location of primary stress was indicated by an accent mark in the visual presentation. This probably made the speaker consciously pay attention to the stress pattern differences between the noun forms (*TRANSplant*) and the verb forms (*transPLANT*), and made him acoustically distinguish those two forms with some exaggeration. As a result, durational contrast between the two stress patterns in Sugahara (2016a) was clearer than that in Sugahara (2011). That is, the listeners were able to discriminate the two stress patterns based on durational cues in Sugahara (2016a) even when pitch was not a reliable cue to discriminate them. This means that there was no room for their knowledge about stress distribution to play a role there.

2.1.3. Predictions for native English listeners

In this section, predictions are made for the native English listeners responses to stimuli used in the current study.[7]

<Group 1: *TRANS*- and *trans*- from *TRANSplant* and *transPLANT*>

The Group 1 stimuli consisted of only the initial syllables of disyllabic words such as *TRANS*- and *trans*- extracted from *TRANSplant* and *transPLANT*. When the truncated stimuli contain substantial acoustic information (pitch, duration, etc.) to distinguish the two stress patterns, native English listeners will rely on it and will be able to correctly identify the original words from which the stimuli are extracted. However, when they lack such acoustic information, they become uncertain about their identity and they will rely on frequency-related knowledge. The following is the list of frequency-related knowledge that may possibly affect native listeners' perception.

Table 1. The list of frequency-related knowledge

		expected effect
(i)	The English lexicon as a whole is dominated by initial main stress (see Section 2.1.1.).	Initial-stress bias
(ii)	Initial main stress is more common than final stress in disyllabic words (see Section 2.1.1.).	
(iii)	There are more nouns in the English vocabulary than verbs, and nouns tend to have stress in an earlier position than verbs (Hammond, 1999).	
(iv)	Prefixes such as *trans*- and *mis*- are less likely to receive main stress (see Section 2.1.1).	Final-stress bias
(v)	There is little difference in lemma frequency between *TRANSplant* and *transPLANT* (see Figure 1, Section 3.1).	No bias towards either direction

The knowledge pertaining to (ii) and (iii) may be part of the knowledge about the lexicon as a whole shown in (i), and therefore (i) to (iii) may be counted as a single set, all of which will lead participants towards an initial stress bias. The knowledge in (iv), however, will lead participants towards the opposite direction. As for (v), the lemma frequency of initial-stress words and that of final-stress words used in this study are almost equal as shown in Figure1, Section 3.1.1. Therefore, the knowledge in (v) will not lead the participants towards either direction.

If those three knowledge factors affect the participants additively and their effect size is equal, native listeners will not be biased towards either of the stress patterns because the factors leading them to the opposite directions cancel each other's effect. If the knowledge factors in (i) to (iii) affect their perception more strongly than the others, their responses will be biased towards initial stress. If the other way round is true, then the outcome will be the opposite. In this way, it is impossible to make a prediction for native English listeners' response patterns based on the frequency-related knowledge shown in Table 1.

The fact that Sugahara (2011) obtained an initial-stress bias using the full forms of the noun-verb pairs implies a possibility that the factors in (i) to (iii) are more influential than the factor in (iv).

<Group 2: *DOmi*- and *DOmina*- from *DOminating*, *domi*- and *domiNA*- from *domiNAtion*>

Secondly, consider the Group 2 stimuli, i.e., the truncated stimuli extracted from four-syllable suffixed words. One group of those four-syllable words have initial stress ending with an inflectional suffix -*ing*, e.g., *DOminating*, and the other have penultimate stress (stem-final stress) ending with a derivational suffix -*ion*, e.g., *domiNAtion*. The base forms of those suffixed words are trisyllabic verbs having initial stress and ending with a morpheme -*ate*, e.g., *DO.mi.nate*. There are two types of truncated stimuli: two-syllable stimuli (*DOmi*- and *domi*-) and three-syllable stimuli (*DOmina*- and *domiNA*-). Since the truncated stimuli extracted from those forms do not provide native English listeners the full segmental and suprasegmental information of their original words, frequency-related knowledge is likely to play a role in their responses. The frequency-related knowledge that need be taken into consideration here is summarized below.

Table 2. The list of frequency-related knowledge

		expected effect
(i)	The English lexicon as a whole is dominated by initial main stress (see Section 2.1.1.).	Initial-stress bias
(ii)	There is little difference between the frequency of the lemmas of the *-ing* forms (e.g., *motivate*) and that of the *-ion* forms (e.g., *motivation*) (see Figure 7, Section 4.1.).	No bias towards either direction
(iii)	The frequency of the *-ing* forms (e.g., *dominating*) is lower than that of the *-ion* forms (e.g., *domination*) (see Figure 6, Section 4.1.).	Stem-final-stress bias

Here, it is almost impossible to make a prediction for native English listeners' response patterns. If the knowledge in (i) affects them the most strongly, they will be biased towards initial stress. If the knowledge in (iii) affects them the most strongly, however, they will be biased towards stem-final stress.

2.2. Japanese

2.2.1. Accent distribution in Japanese

Japanese is a pitch accent language in which lexical prominence is acoustically signaled mostly by a high-low falling pitch (Beckman, 1986; among others). It is also a language in which both the presence or absence and the location of accent are distinctive. Therefore, there is a minimal triplet such as *hasi* ('edge', accentless), *ha˥si* ('chopsticks', accent on the first syllable) and *hasi˥* ('bridge', accent on the final syllable) in Tokyo Japanese, and a minimal pair such as *hasi* ('chopsticks' accentless) vs. *ha˥si* ('bridge', accent on the first syllable) in Kyoto-Osaka Japanese.

A closer look into the Japanese lexicon gives an interesting picture: the location of accent in accented words is strongly skewed to the syllable that contains the antepenultimate mora. Take three-mora accented words for example: Kubozono (2006) has shown that more than a half of the native (Yamato) words and approximately 95% of the Sino-Japanese words and the loanwords have an accent on the antepenultimate mora. Similar figures are true of Osaka Japanese according to Sugahara's (2016a) count of accented words stored in Sugito's (1995) *Osaka/Tokyo akusento onsei ziten* (Pronunciation dictionary of Osaka/Tokyo accent). Tanaka (2009) has also looked into accented loanwords with five morae in both Tokyo and Osaka Japanese, and reports that antepenultimate-mora accent is the most common, i.e., about 40 to 50% of the accented loanwords have accent on their antepenultimate mora.

2.2.2. A bias towards antepenultimate-mora stress?

Having in mind the fact that the Japanese lexicon is dominated by antepenultimate-mora accent, consider the English noun-verb pairs such as *TRANSplant* (noun) and *transPLANT* (verb) used in Sugahara's (2011, 2016a) auditory tasks. The final syllable is a superheavy ending with a consonant cluster, and is therefore divided into three different moras: /a/, /n/, /t/. The nucleus vowel /a/ of the final syllable is the antepenultimate mora, which is expected to be the most favorable prominence position for native Japanese speakers regardless of whether the words are originally of initial stress or of final stress in English. This becomes apparent in loanword adaptation. If *TRANSplant* (noun) and *transPLANT* (verb) are adapted into Japanese as a loanword, both of them come to have accent on the antepenultimate mora vowel as in *t‹o›.ra.n.s‹u›.p‹u›.**ra**⌝.n.t‹o›*, where the vowels in '‹ ›' are epenthetic and '.' separates moras.[8]

Here, one should not confound the antepenultimate mora in Japanese with the word-final syllable in English: not all vowels in word-final syllables in English are antepenultimate moras. Only when word-final syllables are superheavy, their nucleus vowels are antepenultimate. When words end with a heavy or a light syllable such as *hotel* and *banana*, native Japanese speakers are not attracted to their final syllable vowels as the most desirable accent position. Once those words are introduced into Japanese as loanwords, Japanese speakers place accent on the nucleus vowel of their initial syllables as *ho⌝.te.r<u>* and *ba⌝.na.na*, because the initial syllable vowels correspond to their favorite antepenultimate mora.

Let us now consider native Japanese listeners' results obtained in Sugahara's 2011 and 2016a studies. In both cases, Japanese listeners showed strong final bias when they heard trochaic and iambic English words such as *TRANSplant* and *transPLANT*. Sugahara (2016a) interpreted the result as supporting evidence for Japanese listeners relying on the knowledge of antepenultimate-mora accent in Japanese (L1) even in tasks where they judge English (L2) stress locations. The current study, then, asks whether a similar final-stress bias is obtained in tasks with truncated English words.

2.2.3. Predictions for Japanese listeners in this study

<Group 1>

Let us consider what predictions are made for Japanese listeners in this study. As for the Group 1 stimuli (*TRANS*- from *TRANSplant* and *trans*- from *transPLANT*), they are predicted to prefer more final stress than the other two groups of listeners because the final syllable vowel corresponds to the antepenultimate mora.

This, however, does not necessarily mean that Japanese listeners have not acquired at all the frequency-related knowledge of English in Table 1, Section 2.1.3. For example, Ishikawa and Nomura (2008) report that not only native English speakers but also Japanese speakers use probabilistic knowledge about the stress patterns of verbs and nouns when predicting the stress patterns of pseudo English words in a production task. My assumption here is that even if Japanese listeners have acquired frequency-related knowledge about the English vocabulary, their preference for antepenultimate stress is more influential than the frequency-related knowledge of English when perceiving English stress. As a result, they are predicted to be more biased towards final-syllable stress than English native listeners.

<Group 2>

As for the truncated stimuli extracted from Group 2 (*DOmi*- and *DOmina*- from *DOminating* and *domi*- and *domiNA*- from *domiNAtion*), it is predicted that Japanese listeners are more likely to respond that the truncated stimuli are from the *-ion* forms. It is because the *-ion* forms match the stress pattern ideal to them.

In the forced identification task of this study with those stimuli, every time they hear one of the stimuli, they are visually given two options from which they are required to choose as what they believe to be its original word from. Those options are always an *-ing* form with an accent diacritic on the initial syllable vowel, e.g., *dóminating*, and an *-ion* form with an accent diacritic on the penultimate syllable vowel, e.g., *dominátion* (see Figure 12, Section 4.2). The representation *dóminating* has stress on its initial syllable, which is too far from the ideal stress position, i.e., the antepenultimate mora. On the other hand, the representation *dominátion* has stress on the ideal position, i.e, on the antepenultimate mora as shown in (1a). When the word is adapted into Japanese as a loanword, it also carries an accent on the same antepenultimate vowel as in (1b).

(1)

a. *do. mi. n*[éɪ]. [ʃn̩] (English)

b. *do. mi. ne˥e. ʃon* (Japanese loanword)

Given this, it is highly likely that Japanese listeners accommodate the stimuli auditorily presented to them to the -*ion* forms, e.g., *dominátion*.

2.3. Seoul Korean

Seoul Korean is a language with neither a lexical stress nor a lexical accent, and pitch does not play a distinctive role. Although it has tonal melodies, they are the properties of accentual phrases and intonational phrases (Jun, 1996, 1998, 2005, 2006). According to Jun, long accentual phrases, i.e., those with four or more than four syllables, are associated with a sequence of LHLH or HHLH tones unless they are intonational phrase-final[9], and whether or not they start with LH or HH depends on the segmental content of their initial segment. If the initial segment is a 'lenis' obstruent or a sonorant, the accentual phrase begins with a LH tone. If it is an 'aspirated' or 'fortis (tense)' obstruent, however, an HH tone appears instead (Jun, 1998). When it comes to shorter accentual phrases, medial tones undergo 'undershoot' and they come to have tonal melodies such as L(HL)H, L(H)LH, LH(L)H, H(HL)H, H(H)LH, HH(L)H, where the tones in the parenthesis are the ones that are undershot (Jun, 2005, 2006). The final H tone, however, is always at the final syllable of an accentual phrase unless the accentual phrase is intonational phrase-final. Jun (1996) further states that the accentual-phrase final H tone is 'salient' and is realized with higher pitch than the H tones that may appear earlier in an accentual phrase. It is not clear in what sense Jun (1996) used the word 'salient', whether it is simply higher than the other H tone in F0 or if it even has perceptual salience. The salience, whether phonetic or perceptual, does not have a root in the phonemic representation of the lexicon. Because the final H tone is the property of an accentual phrase, the final syllable of the same word does not always bear the H tone: whether the H tone appears or not depends solely on which position it occupies in an accentual phrase. Accentual phrase formation in Seoul Korean varies depending on speech rate, phonological weight, information structure, semantic weight and morpho-syntactic structure (Jun, 1996). For example, a four-word sentence in (2) may consist of four separate accentual phrases if it is produced with a slow speech rate while it may have only two accentual phrases if it is produced with a fast rate (Jun, 1996).

(2)	/ igən	adʒu	tʃoɨn	kɨɾimija /		
	this	very	good	a picture-be		
a.	{L H}	{L H}	{L H}	{LH	IntPh-final tone}	slow rate
b.	{L H}	{LH			IntPh-final tone}	fast rate

(from Jun, 1996: 158)

The third lexical word *tʃoɨn* 'good' in (2a) forms an independent accentual phrase and its final syllable bears an H tone. The same syllable in (2b), however, does not carry an H tone because it is phrase-medial. Furthermore, when accentual phrase-final syllables are at an intonational phrase-final position in a declarative utterance, it is no longer associated with an H tone but is associated with an L boundary tone. In this way, the presence or absence of an H tone in word-final syllables is purely post-lexical and does not play a distinctive role in the phonemic representation.

Given this, it is imaginable that native Seoul Korean speakers set their ‘stress parameter’, the term originally proposed by Peperkamp & Dupoux (2002), such that lexical stress/accent information is not encoded in their phonological representation because it is not useful in their native language, and the parameter even affects their L2 perception. That is, it is possible that Seoul Korean learners of English are less fastidious about the location of lexical stress than Japanese learners of English when asked to judge which syllable is stressed in English words.

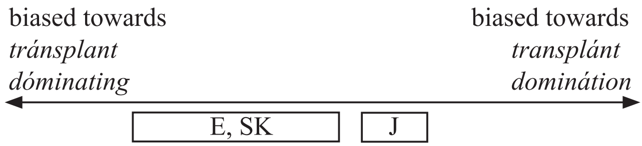
As already introduced in Section 1, Guion (2005) has found that in her production and perception study using pseudo English words that late Seoul Korean-English bilinguals did not distinguish the stress pattern of nouns and that of verbs as much as native English speakers and early Seoul Korean-English bilinguals did. Guion states that it is because their L1 lacks lexical stress and they pay less attention to stress pattern differences between nouns and verbs when learning English vocabulary.

If so, it is expected that the Seoul Korean listeners in this study are less influenced by the statistical distribution of English stress than native English listeners. In addition, they do not have any preferred lexical stress/accent position in L1. Given this, it is at least predicted that Seoul Korean listeners’ responses are less biased towards final stress (Group 1) and are less biased towards *-ion* forms (Group 2) than Japanese listeners’. However, no clear prediction is made as to whether they would give responses different from native English listeners, because the only prediction made for native English listeners is that they will be less biased towards final stress than Japanese listeners (Section 2.1.3 and Section 2.2.3).

2.4. Summary of predictions

For both Group 1 and Group 2, the only prediction made in this section is that Japanese listeners are more biased towards final-syllable stress in the task with the Group 1 stimuli and they are more biased towards the -*ion* forms in the task with the Group 2 stimuli than native English listeners and Seoul Korean listeners. No prediction is made for the relation between native English listeners and Seoul Korean listeners. This is summarized below.

(3) Predictions



3. The identification task with the truncated stimuli ‘Group 1’

3.1. Method

3.1.1. Stimuli

The truncated stimuli considered here consist of initial syllables extracted from two-syllable noun-verb minimal pairs such as *TRANSplant* (noun) and *transPLANT* (verb), which are shown in (4). Their full-word counterparts were used as stimuli in the task reported in Sugahara (2016a).

(4) Truncated stimuli ‘Group 1’

a. *IM*- from *IMport*, *im*- from *imPORT*
b. *IN*- from *INsult*, *in*- from *inSULT*
c. *MIS*- from *MISprint*, *mis*- from *misPRINT*
d. *RE*- from *REtake*, *re*- from *reTAKE*
e. *TRANS*- from *TRANSplant*, *trans*- from *transPLANT*

The log lemma frequencies[10] of those words were obtained from the CELEX database (Baayen, Pipenbrock & Gulikens, 1995) and are summarized in Figure 1.[11] The frequency information is considered here because there is a possibility that listeners gravitate to more frequent forms when asked to identify what they have heard.

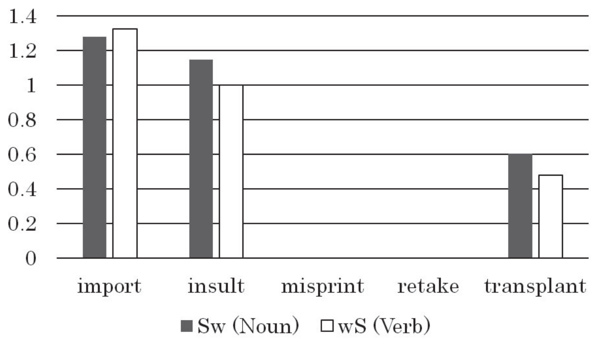


Figure 1. The log lemma frequency of the stimuli words obtained from CELEX. The dark gray bars are of trochaic nouns (e.g., *TRANSplant*) and the white bars are of iambic verbs (e.g., *transPLANT*).

The mean log frequency of the trochaic noun forms (*TRANSplant*) turned out to be 0.61 and that of the iambic verb forms (*transPLANT*) was 0.56, which are very close. Given this, it is safe to say that the lemma frequency of the words used in this task does not affect participants' responses.

The stimuli were created through the following procedure. Their full form counterparts, which were used in Sugahara (2016a), were first produced by a male American English speaker in his mid-twenties. The full words were embedded in a sentential frame "*I wanted to say* ___" when they started with a consonant and in "*I said* ___" when they started with a vowel or a liquid /r/. The location of primary stress was indicated by an accent mark '´' on the vowel bearing main stress. Those sentences were recorded onto a Marantz Solid State Recorder PMD671 (44.1 KHz, 16 bits), using a Countryman ISOMAX Headset Microphone in a sound-attenuated room. The target words (full words) were segmented from the sentence frames, and the initial syllables were further segmented from the full words using Praat (Boersma & Weenink, 2013).

Two kinds of F0 conditions were, then, prepared: <Natural> and <90Hz>. The former consisted of stimuli without any pitch manipulation. The latter underwent pitch manipulation and the F0 of the vowel interval was made into 90Hz.

Acoustic properties of the <Natural> stimuli are summarized in Table 3. From the table, it is clear that the initial syllables of the trochaic noun forms, e.g., *TRANS*-, had greater values than those of the iambic verb forms, e.g., *trans*-, in three acoustic dimensions, i.e., F0, duration, and overall intensity. The high-vowel syllables (*im*-, *in*-, *mis*-, *re*-) in the noun (trochaic) forms had lower F1 and higher F2 than those in the verb (iambic) forms, which means that the former were produced with more jaw closure and more tongue advancement. For the low-vowel syllable (*trans*-), its noun form had a slightly higher F1 value (more jaw opening) and a lower F2 (less tongue advancement) than the verb form. Those acoustic measures show that although the truncated stimuli contained less information to disambiguate the two stress patterns than the full-word stimuli used in Sugahara (2016a), they still had some acoustic contrast between the two stress patterns.

Table 3. Acoustic properties of the truncated <Natural> stimuli

Initial syllable	Category	Stress	F0 averaged across the vowel period (Hz)	Duration (ms)	Overall intensity (dB)	Midpoint F1 (Hz)	Midpoint F2 (Hz)
im(port)	N	primary	121	145	72	500	2517
	V	secondary	87	131	63	597	2071
in(sult)	N	primary	106	187	67	457	2439
	V	secondary	86	156	64	543	2209
mis(print)	N	primary	98	220	69	470	1843
	V	secondary	85	182	62	500	1036
re(take)	N	primary	111	151	63	354	1702
	V	secondary	88	130	59	369	1629
trans(plant)	N	primary	100	329	67	659	1464
	V	secondary	87	264	63	644	1644

3.1.2. Participants

Twenty-one native English listeners (eight males and thirteen females), thirty native Japanese listeners (twelve males and eighteen females), and twenty-seven Seoul Korean listeners (four males and twenty-three females) participated in the identification task. The native English speakers were exchange students studying at Doshisha for a year or a semester, or international students studying at Doshisha University Center for Japanese Language and Culture for a year[12], and all of them had the experience of learning Japanese for one to three years. Fourteen were from the US, three were from England, two were from Canada, and the other two were from Australia.[13] The native Japanese speakers were full-time students at Doshisha University or Kyoto University and most of them were from the

Kansai area speaking Kansai Japanese as their native dialect. They had been learning English as a second language for more than six years prior to the experiment. They were asked to report their TOEFL ITP, TOEFL iBT or TOEIC scores if they had taken any of those English proficiency tests in advance. Twenty-eight of them revealed their scores. Their scores were converted into CEFR (Common European Framework of Reference for Languages), and their CEFR proficiency were the following: three of them were at A2 (waystage), fourteen of them were at B1 (threshold) and eleven of them were at B2 (vantage). All of the Seoul Korean listeners were university students, too. Nineteen of the Seoul Korean listeners were full-time students at Ewha Womans University in Seoul and they were visiting Doshisha to take part in a two-week spring program when they participated in the experiment. The rest of the Seoul Korean listeners were international students at Doshisha University Center for Japanese Language and Culture. All of the Korean listeners had studied English and Japanese as second languages (English for more than six years and Japanese for one to nine years) prior to the experiment. Although they were asked to report their English proficiency test scores, only eight of them did so. Among them, one was at B1 and seven were at B2. The rest of the Koreans were also able to communicate in English without any difficulty and all the instructions and written information were given to them using English. None of the participants had reported hearing disorders, and all of them were paid for their participation. As already stated in Sugahara (2016a), they were given a chance to be reminded of the stress pattern differences between the noun forms and the verb forms listed in (4) before the perception task. For more details about this, see Sugahara (2016a: 93).

3.1.3. The forced choice identification task

The identification task was carried out via SuperLab Version 4.5 installed on MacBook Air with OS X Version 10.7.4 (also see Sugahara (2016a) for more details). The <Natural> and the <90 Hz> truncated stimuli were presented to participants in separate blocks in the order shown in (5). The order of stimuli presentation within each block was randomized for each participant, and each stimulus was presented only once to each participant.

(5) Presentation order of blocks
(F) → (O) → (O) → <Natural> → (F) → (O) → (O) → <90Hz> → (F)

F = blocks with full word stimuli (noun-verb minimal pairs) used in Sugahara (2016a)
O = blocks with other truncated words

The participants listened to the stimuli through SONY dynamic stereo headphones connected to MacBook Air (also see Sugahara (2016a) for more details). The letter strings of the noun-verb pair to which the stimulus belonged were presented on a computer monitor whenever each truncated stimulus was played to a participant. The letter sequence that corresponded to the initial syllable of a noun (trochaic) stimulus was colored in yellow and those which corresponded to the initial syllable of a verb (iambic) stimulus was colored in blue. Truncated parts of the stimuli were colored in light gray. The letter sequence of a trochaic form was shown at the top and that of an iambic form was shown below it. The location of primary stress was indicated by a diacritical mark ‘′’, which was placed above the letter corresponding to the nucleus of a syllable with primary stress as shown in Figure 2. The participants pressed the yellow button on the computer keyboard if they thought the stimulus they had heard was trochaic and pressed the blue button if they thought it was iambic. They only used their index finger of their dominant hand when carrying out the task. (Also see Sugahara (2016a) for relevant and more detailed information).

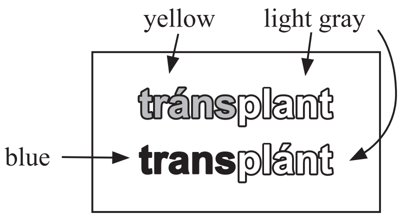


Figure 2. A sample of the visual presentation of response options on the computer monitor

3.2. Results and discussion

The number of responses obtained from each participant was 10 (5 pairs × 2 stress patterns), and 1,560 responses (10 responses × 2 blocks × 78 participants) were obtained in total. Then, the correct response rate of each language group was calculated for each stress and pitch condition, which is summarized in Figure 3.

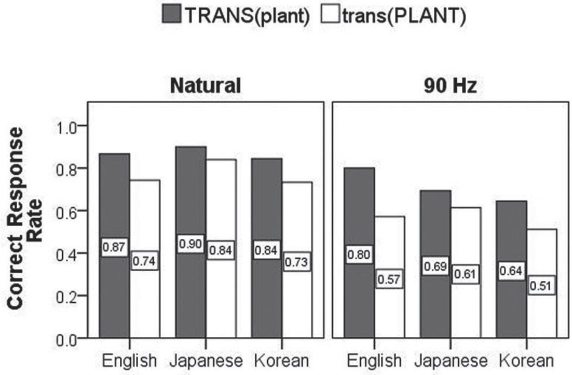


Figure 3. The rate of correct responses in the Natural and the 90 Hz condition.

Overall, the rates of correct responses were greater in the Natural condition than in the 90 Hz condition regardless of the stress patterns of the stimuli and the participants' native languages. It is not surprising given that the Natural condition provided listeners with F0 cues to disambiguate the two stress patterns. Another thing is that the correct response rates of the stimuli originally produced as trochaic were greater than those of the stimuli originally produced as iambic regardless of the difference in native languages. That is, all language groups showed a trochaic bias rather than an iambic bias. Nonetheless, the Japanese listeners' trochaic bias was the weakest of all.

To see if the difference between the Japanese group and the other two language groups is real, a bias measure '*c*' (Macmillan & Creelman, 1990, 2005) was obtained for each listener in each stress and each F0 condition. Since the main goal of this study is to examine if the Japanese listeners are biased towards final-stress (antepenultimate-mora-stress) even when presented with truncated stimuli, *c* values considered here were calculated to show the magnitude of a final-stress bias, though all the three language groups did not actually show such a bias here.[14] Figure 4 summarizes the mean *c* values of each language group averaged across listeners in each F0 condition.

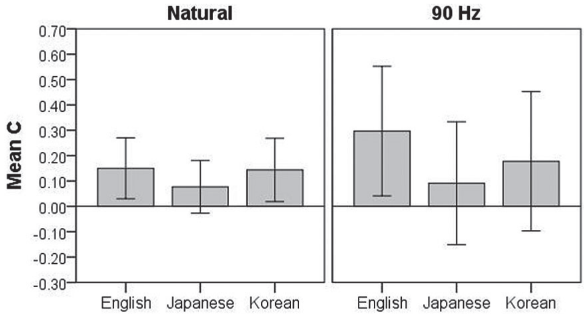


Figure 4. The mean *c* values of each language group averaged across listeners in each F0 condition. Error bars show 95% confidence intervals.

The distance between *c* and zero indicates the magnitude of a bias: the farther apart from zero in the negative range, the more biased towards final stress, and the farther apart from zero in the positive range, the more biased towards initial stress. If *c* is zero, then there is no bias towards either direction. Since all language groups were biased towards trochaic stress in this case, their mean *c* values were all in the positive range as shown in Figure 4. Nonetheless, the mean *c* values of the Japanese listeners were the closest to zero, and their 95 % confidence intervals in both the Natural and the 90 Hz condition included zero. That is, although their mean *c* values were indeed in the positive range, i.e., in the direction of an initial-stress bias, it was too weak to conclude that there was any bias. The mean *c* value of the English listeners and that of the Seoul Korean listeners in the Natural condition were greater in the positive range than that of the Japanese listeners, and their 95 % confidence intervals did not include zero, which means that the English and the Seoul Korean listeners were biased towards trochaic stress. When it comes to the 90 Hz condition, it is only the English listeners who showed a clear trochaic bias because it is only their 95 % confidence interval that did not include zero. In summary, the Japanese listeners were the least biased and the English listeners were the most biased towards initial stress.

The difference among the language groups, however, was not statistically significant according to ANOVA (univariate general linear model) in which *c* was a dependent variable and 'language groups' and 'F0 conditions' were the fixed factors: There was no significant difference between the two F0 conditions, and there was no significant interaction between the language factor and the F0 factor, either. That is, although the native English listeners showed the strongest and the Japanese listeners showed the weakest trochaic bias, it was not backed up by the results of statistical analysis.

The results here are in contrast with Sugahara's (2016a) observation that the Japanese and the Seoul Korean listeners showed a final stress bias when they were provided with full-word stimuli such as *TRANSplant* and *trasnPLANT* in an environment where pitch was not a reliable cue.[15] One possible account is the following. First of all, the final syllables of the full-word stimuli in Sugahara (2016a) were longer than their initial syllables as summarized in Table 2. For example, the durational ratio between the initial and the final syllable in the trochaic noun form *IMport* is 1:3.5. The only pair whose final syllables were less than two times longer than their initial syllables was the pair of *transplant*.

Table 4. The durational properties of syllables in the full-word stimuli used in the task reported in Sugahara (2016a)

Initial syllable	Category	Stress Pattern	Syllable 1 Duration(ms)	Syllable 2 Duration(ms)	ratio between Syllable 1 and Syllable 2
import	N	trochaic	145	508	1:3.5
	V	iambic	131	575	1:4.4
insult	N	trochaic	187	501	1:2.7
	V	iambic	156	585	1:3.8
misprint	N	trochaic	220	454	1:2.1
	V	iambic	182	524	1:2.8
retake	N	trochaic	151	431	1:2.9
	V	iambic	130	456	1:3.5
transplant	N	trochaic	329	481	1:1.5
	V	iambic	264	506	1:1.9

Given this, it is possible that both the Japanese and the Seoul Korean listeners paid more attention to the relative durational difference between the initial and the final syllables than the native English listeners, and they might have judged that the longer final syllables had been more prominent than the shorter initial syllables. A closer investigation into the data considered in Sugahara (2016a) seems to support this speculation. Although Sugahara (2016a) only showed correct response rates for the full-word stimuli having all of the word-pairs lumped together, Figure 5 shows those of each word pair separately.

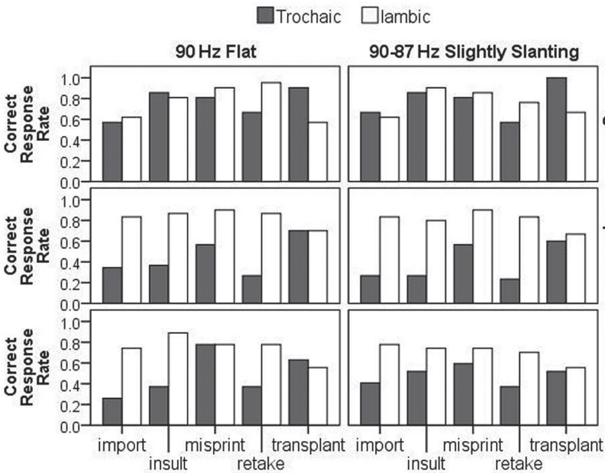


Figure 5. The correct response rates obtained for the full-word stimuli in Sugahara (2016a). '90 Hz Flat' means that the pitch contours of the stimuli were made into flat at 90 Hz, and '90-87 Hz Slightly Slanting' means that the pitch contours of the stimuli were made into a slightly slanting slope declining from 90 Hz to 87 Hz.

Figure 5 shows that listeners were paying attention to durational difference between the initial and the final syllables. The evidence comes from the pair of *transplant* which behaves differently from the other pairs. For the Japanese listeners, it is only the pair of the *transplant* that obtained no bias towards either direction. A similar picture was also basically true of the Seoul Korean listeners. When it comes to the English listeners, only the pair of *transplant* obtained a strong initial-stress bias. This peculiar behavior of the *transplant* pair is explained if the durational difference between the initial and the final syllable is taken into consideration: *transplant* was the only pair whose initial and final syllables had relatively similar duration. The two syllables having similar duration might have made the participants hear less prominence on the final syllables, which might have resulted in no final stress bias or a bias towards initial stress. In this way, the relative length factor of the initial and the final syllables might have affected the results of the task reported in Sugahara (2016a).

In the current task, the participants were not able to hear the relative durational difference between the initial and the final syllables. Therefore the Japanese and the Seoul Korean listeners were not biased towards final stress as they did in Sugahara (2016a). Since the relative durational difference between the initial and the last syllable in the same word is not available, one may speculate that the current task allowed more genuine effect of the stress/accent systems of listeners' L1 to emerge. However, in this task, no significant difference was obtained between the response pattern of Japanese listeners and that of the other two language groups. In summary, no evidence was obtained to support the hypothesis that the Japanese listeners are affected by the antepenultimate stress rule in their L1 when perceiving English stress.

4. The identification task with truncated stimuli 'Group 2'

4.1. Stimuli and predictions

The second group of truncated stimuli comes from four-syllable suffixed words sharing the same verb stems: the *-ing* forms such as *DOminating* and the *-ion* forms such as *domiNAtion*. The complete list of those words used in the experiment are shown below.

(6)

a. *ACtivating,* *actiVAtion*

b. *Agitating,* *agiTAtion*

c. *ALlocating,* *alloCAtion*

d. *CALculating,* *calcuLAtion*
e. *CAPtivating,* *captiVAtion*
f. *CElebrating,* *celebRAtion*
g. *COMplicating,* *compliCAtion*
h. *CONcentrating,* *concentRAtion*
i. *CONjugating,* *conjuGAtion*
j. *DEdicating,* *dediCAtion*
k. *DOminating,* *domiNAtion*
l. *Educating,* *eduCAtion*
m. *EStimating* *estiMAtion*
n. *INdicating,* *indiCAtion*
o. *MEdicating,* *mediCAtion*
p. *MItigating,* *mitiGAtion*
q. *MOtivating,* *motiVAtion*
r. *NAvigating,* *naviGAtion*
s. *PROpagating,* *propaGAtion*
t. *PROsecuting,* *proseCUtion*
u. *TERminating,* *termiNAtion*

It is also necessary to inspect the frequency of each word to see if word frequency affected listeners responses. Frequency information was obtained from the CELEX database (Baayen, Piepenbrock & Gulikers, 1995) and is summarized in Figure 6.

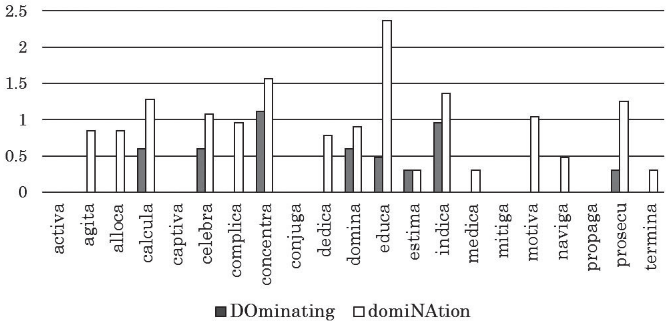


Figure 6. The log frequency obtained from CELEX. The dark gray bars are of the -*ing* forms (e.g. *DOminating*) and the white bars are of the -*ion* forms (e.g., *domiNAtion*).

The mean log frequency of the -*ing* forms is 0.24 while that of the -*ion* forms is 0.75. The log value of 0.24 means that there are about 1.8 occurrences and that of 0.75 means that there are about 5.6 occurrences, which means that the -*ion* forms are about three times more likely to occur. This frequency difference may affect the result of the experiment, and the frequency factor was taken into consideration in the analysis presented in Section 4.3.

The log frequency of verb lemmas, i.e., the frequencies of all inflected forms combined together, is also shown in Figure 7. The mean verb lemma frequency is 0.86 (7.3 occurrences) and the mean frequency of the -*ion* forms is 0.74 (5.5 occurrences), which means that the total number of occurrences of the former is 1.3 times higher than that of the latter. The difference, however, is not as obvious as the difference between the inflected forms -*ing* and the derived forms with -*ion* shown in Figure 6, which I assume not to have affected the responses of the participants.

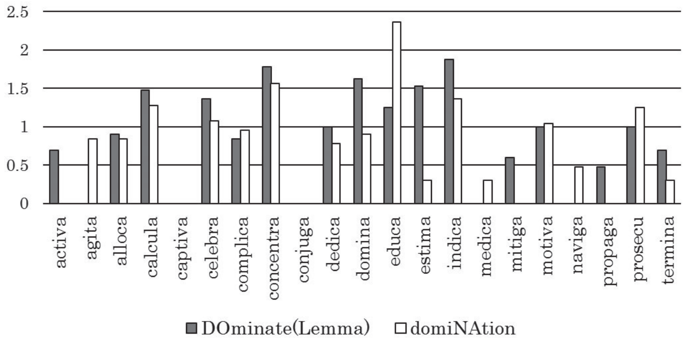


Figure 7. The log frequency obtained from CELEX. The dark gray bars are of the verb lemmas (e.g. *DOminate*) and the white bars are of the *-ion* forms (e.g., *domiNAtion*).

4.2. Method

The words listed in (6) were produced by the same native English speaker who also produced the Group 1 stimuli. The recording procedure was the same as that of the Group 1 task. The only difference was that although the primary stress locations of the Group 1 words were indicated by an accent mark '´' when they were visually presented to the speaker, no such marking was employed for the words in (6). It is because for the native speaker the stress patterns of the words in (6) were obvious from their suffix endings.

Two kinds of truncated words were created using Praat: those with initial two syllables such as *DOmi*- from *DOminating* and *domi*- from *domiNAtion*, and those with three syllables such as *DOmina*- from *DOminating* and *domiNA*- from *domiNAtion*. Acoustic properties of those stimuli are summarized in Tables 5 to 8 and Appendices 1 to 3.

As shown in Table 5 and Appendix 1, the initial vowels and syllables of the *-ing* forms were longer than those of the *-ion* forms. It is also true that the second vowels and syllables of the former were slightly longer than those of the latter. Their third vowels and syllables, however, displayed an opposite relation: the -*ion* forms were longer than the -*ing* forms. That is, the durational differences between the -*ing* and the -*ion* forms may serve as disambiguating cues if listeners are sensitive to them. Table 6 shows the mean values of F0 averaged across each vowel period and those of overall intensity: the initial and the second vowels of the -*ing* forms have greater mean values than those of the -*ion* forms while the other way round is true for their third vowels (also see Appendix 2). Those may also serve as disambiguating cues. Formant values are shown separately for the initial vowels and for the final vowels in Table 7 and Table 8 respectively (also see Appendix 3). The magnitude of differences in F1 mean values between the -*ing* and the -*ion* forms was small. The only initial vowels that differed relatively clearly (about 40 Hz or more) between the two forms were /i/ and /æ/. As for the third vowels, there were no such differences. The magnitude of differences in mean F2 values was small, too, and the only case where a relatively clear difference (about 100 Hz) existed was the third vowel /ju/ in *prosecuting* and *prosecution*. Given this, formant values are predicted to be less reliable cues for listeners to distinguish the two stress patterns.

Table 5. Duration (ms)

		Vowel1	Vowel2	Vowel3	Syllable1	Syllable2	Syllable3
-*ing*	Mean	99	53	126	173	107	205
	SD	23.9	12.9	12.5	48.6	24.5	27.7
-*ion*	Mean	81	49	132	146	103	223
	SD	18.7	14.3	13.4	47.9	23.5	28.8

Table 6. F0 (averaged across a vowel period) and overall intensity

		Vowel1		Vowel2		Vowel3	
		F0 (Hz)	Intensity (dB)	F0 (Hz)	Intensity (dB)	F0 (Hz)	Intensity (dB)
-ing	Mean	100	64	93	60	79	58
	SD	3.9	2.5	4.8	3.2	2.8	2.5
-ion	Mean	92	62	88	59	90	61
	SD	4.2	2.7	3.5	3.1	3.3	2.4

Table 7. V1 Midpoint F1 and F2 (Hz)

		Front						Non Front					
		ɪ n = 2		ɛ n = 5		æ n = 6		ɝ n = 1		oʊ n = 1		ɑ n = 6	
		F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
-ing	Mean	492	2300	583	1698	771	1591	679	1628	525	1305	686	1155
	SD	51.6	639	61.2	171.8	46.8	198	.	.	.	.	21.5	45.8
-ion	Mean	532	2368	564	1666	719	1504	691	1498	540	1128	690	1182
	SD	62.2	685.2	64.7	183.9	45.7	153	.	.	.	.	18.6	54.1

Table 8. V3 Midpoint F1 and F2 (Hz)

		eɪ n = 20		ju n = 1	
		F1	F2	F1	F2
-ing	Mean	474	1986	380	2031
	SD	30.6	140	.	.
-ion	Mean	458	2013	357	1930
	SD	34.8	153	.	.

The two-syllable stimuli e.g., *DOmi*- and *domi*-, had only one F0 condition <Natural>, whereas the three-syllable stimuli, e.g., *DOmina*- and *domiNA*-, had two F0 conditions: <Natural> and <Slightly Slanting>. This is summarized in Table 9.

Table 9. Truncated stimuli and F0 conditions

truncated stimuli	**F0 conditions**
Two syllables *DOmi*-, *domi*-	<Natural>
Three syllables *DOmina*-, *domiNA*-	<Natural> & <Slightly Slanting>

The <Slightly Slanting> stimuli had the same F0 slope regardless of whether they were extracted from the initial-stress *-ing* forms or from the stem-final-stress *-ion* forms. To create the <Slightly Slanting> stimuli, the left edge of their initial syllable's voicing period was set at 90 Hz and the right edge of their third syllable's voicing period was set at 85 Hz, and those two points were connected with a straight F0 slope using the pitch manipulation function of Praat. In Figures 8 to 11, example pitch contours are shown.

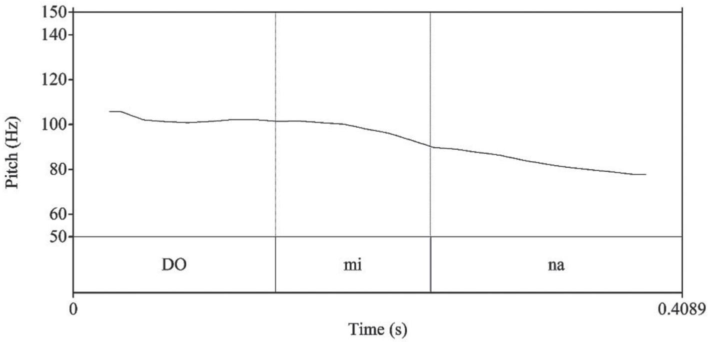


Figure 8. A three-syllable stimulus *DOmina-*: <Natural>

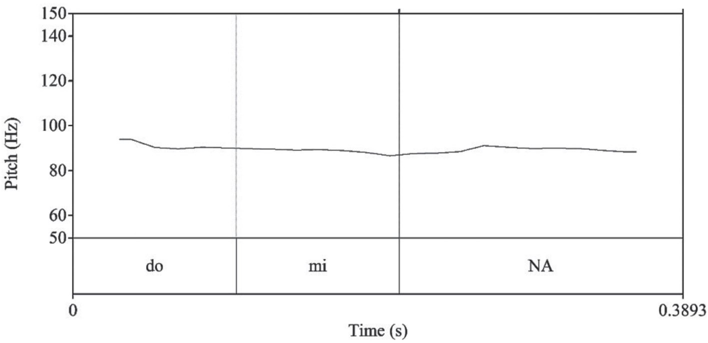


Figure 9. A three-syllable stimulus *domiNA-*: <Natural>

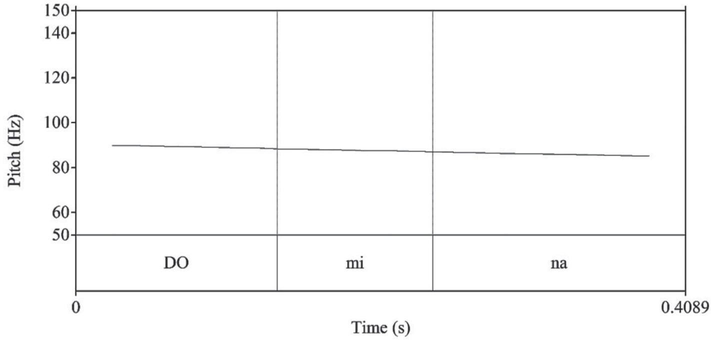


Figure 10. A three-syllable stimulus *DOmina*-: <Slightly Slanting> 90 to 85 Hz

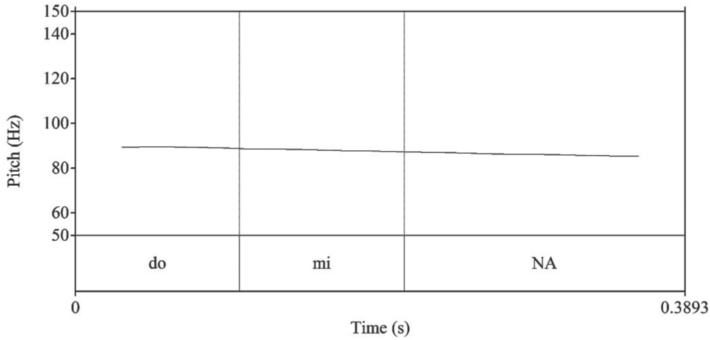


Figure 11. A three-syllable stimulus *domiNA*-: <Slightly Slanting> 90 to 85 Hz

Participants and an experimental procedure were the same as those of the task with the Group1 stimuli presented in Section 3.1. The following is the order of block presentation, and within each block stimuli presentation was randomized for each participant. Each stimulus was presented only once to each participant.

(7) Presentation order of blocks

(F) → <three-syllable: Natural> → <two-syllable: Natural> → (O) → (F) → <three-syllable: Slightly Slanting> → (O) → (O) → (F)

F = blocks with full word stimuli (noun-verb minimal pairs) used in Sugahara (2016a)

O = blocks with other truncated words

An example of the visual presentations with the two options (the *-ing* form and the *-ion* form) on the computer monitor is shown in Figure 12.

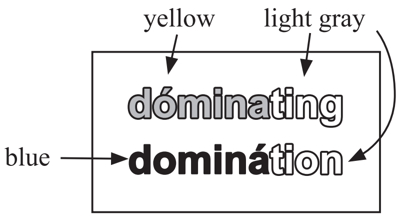


Figure 12. A sample of the visual presentation of options on the computer monitor

4.3. Results

The number of responses obtained from each participant was 42 (21 pairs × 2 stress patterns), and 9,828 responses (42 responses × 3 blocks × 78 participants) were obtained in total. Word pairs were divided into two groups: the pairs where the frequency of the *-ion* forms dominated the *-ing* forms ([*-ing* < *-ion*]: *agita-*, *alloca-*, *calcula-*, *celebra-*, *complica-*, *concentra-*, *dedica-*, *domina-*, *educa-*, *indica-*, *medica-*, *motiva-*, *naviga-*, *procsecu-*, *termina-*) and the pairs where the *-ion* and the *-ing* forms were equal in frequency ([*-ing* = *-ion*]: *activa-*, *captiva-*, *conjuga-*, *estima-*, *mitiga*-, *propaga*-). Then, the correct response rates of each language group was calculated for each block, separately for the [-*ing* < -*ion*] pairs and the [-*ing* = -*ion*] pairs. This is summarized in Figure 13.

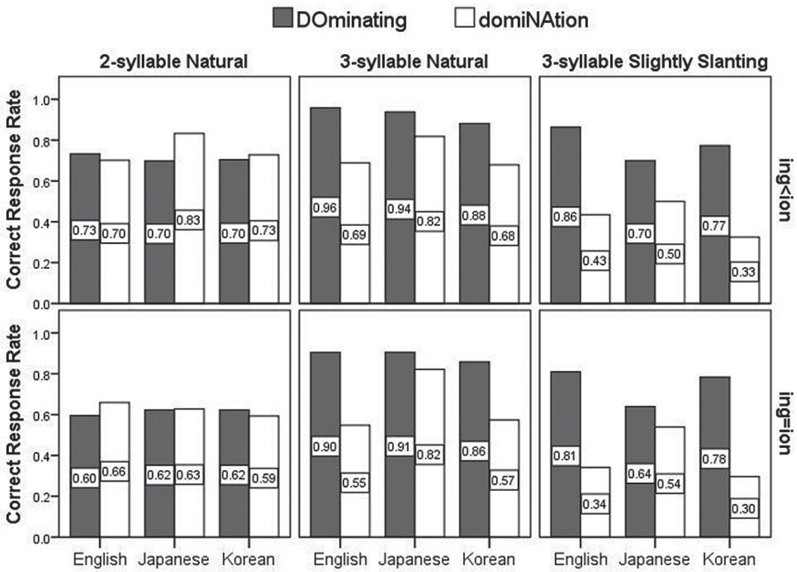


Figure 13. The rate of correct responses in each block

C values were obtained for each speaker in each block separately for the two frequency groups. The calculation method adopted here was the same as the one in Section 3.2.[16] Since this study was originally started to see if the Japanese listeners are biased towards antepenultimate-mora-stress (stem-final-syllable stress), *c* values considered here were calculated to show the magnitude of a bias towards stem-final-syllable stress. As already explained in Section 3.2, the farther apart from zero the *c* values are in the negative range, the more biased towards stem-final stress, and the farther apart from zero in the positive range, the more biased towards initial primary stress. If they are at zero, there is no bias towards either direction.

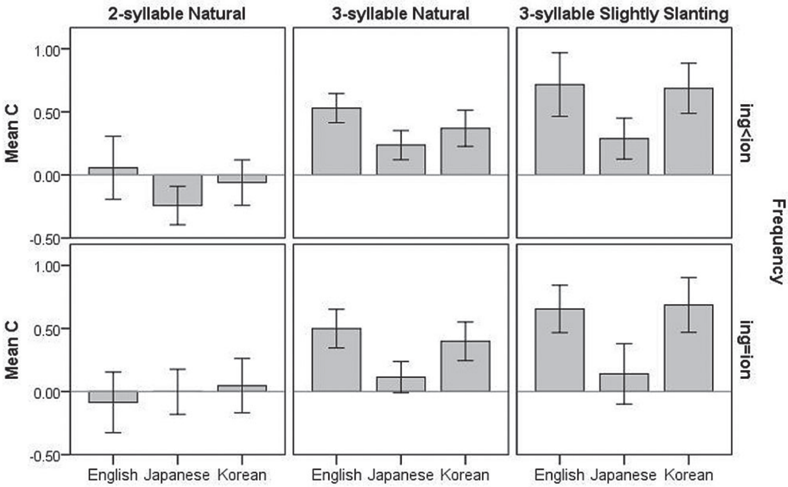


Figure 14. The mean *c* values of each language group averaged across listeners for each block. Error bars show 95% confidence intervals.

There seems to be no or little effect of the frequency factor: both the [-*ing* < -*ion*] case and the [-*ing* = -*ion*] case showed the same response patterns. In the <Two-Syllable Natural> block, the three language groups equally gave neutral responses. The only exception was the Japanese listeners' responses in the [-*ing* < -*ion*] case, where they showed some bias towards *domiNAtion*. However, the 95% CI of their mean *c* value there overlaps with that of the other two language groups, which means that the difference is trivial. In the <Three-Syllable> blocks, regardless of the difference in the frequency factor, all language groups showed a bias towards *DOminating*. The Japanese listeners' bias was the weakest, however. As seen in Figure 13, the Japanese listeners' correct responses to *DOmina-* and those to *domiNA-* exceeded 80% in the <Three-Syllable Natural> block, meaning that they were almost always able to distinguish the two groups of stimuli. Although their correct response rates radically declined in the <Three-Syllable Slightly Slanting> block, the Japanese listeners' responses were relatively neutral compared to the other two language groups who were strongly biased towards *DOminating*.

To see if the difference in the mean *c* values between the Japanese listeners and the other two language groups in the three-syllable blocks is real and also to clarify that the frequency factor did not affect the participants responses in all blocks, a two-way ANOVA was conducted separately for each of the three blocks. In the analyses, *c* was a dependent variable and 'language groups' (English, Japanese, Seoul Korean) and 'frequency conditions'([-*ing* < -*ion*] and [-*ing* = -*ion*]) were fixed factors. Because this is the case of multiple comparisons, i.e., three comparisons were carried out, the level of significance (α) was adjusted to .017 following Bonferroni's correction procedure, dividing the ordinarily adopted level of significance .05 by three.

For the <Two-Syllable Natural> block, both the main effect of language groups and that of frequency conditions were not significant, and there was no interaction between the two factors, either. For the <Three-Syllable Natural> block, only the main effect of language groups was significant (F (2, 150) = 13.42, $p < .001$), and no significant interaction between the two factors was present. For the <Three-Syllable Slightly Slanting> block, too, only the main effect of language groups was significant (F (2, 150) = 14.88, $p < .001$), and no significant interaction between the two factors was present. Tukey-Kramer post-hoc comparisons were carried out for the <Three-Syllable Natural> and the <Three-Syllable Slightly Slanting> blocks where there was a significant effect of the language groups. As it turned out, for both the <Three-Syllable Natural> and the <Three-Syllable Slightly Slanting> blocks, the Japanese listeners were significantly different from the other two groups whereas the difference between the English and the Seoul Korean listeners was not significant (<Three-Syllable Natural>: Japanese vs. English: difference mean = 0.34, $p < .001$, and Japanese vs. Seoul Korean: difference mean = 0.21, $p = .003$; <Three-Syllable Slightly Slanting>: Japanese vs. English: difference mean = 0.47, $p < .001$, and Japanese vs. Seoul Korean: difference mean = 0.47, $p = .002$). Those results are summarized in the following list, and also visually presented in (8).

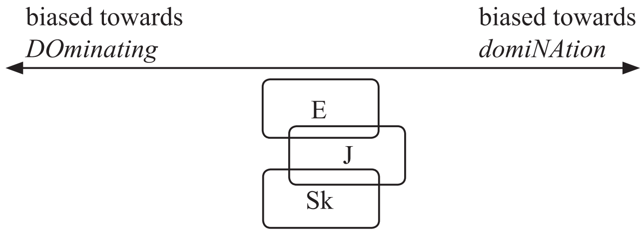
(i) There was no effect of frequency difference between the -*ing* and the -*ion* forms.

(ii) There was a difference between the two-syllable block (*domi*-) and the three-syllable blocks (*domina*-). All three language groups were neutral between the -*ing* and the -*ion* forms in the two-syllable block whereas they were all biased towards the -*ing* forms in the three-syllable block.

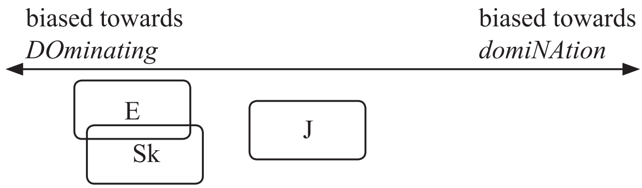
(iii) The Japanese listeners' bias towards the -*ing* forms in the three-syllable block, however, was significantly weaker than the English listeners' and the Seoul Korean listeners' bias. There was no difference between the English listeners and the Seoul Korean listeners.

(8) Results from Group 2

a. Two-syllable stimuli



b. Three-syllable stimuli



4.4. Discussion

Let us consider what the results mean in the light of the hypothesis tested in this study. Although the task with the two-syllable stimuli did not yield the result predicted by the hypothesis, the task with the three-syllable stimuli did yield it. That is, the Japanese listeners were more inclined to prefer *domiNAtion* than the English and the Seoul Korean listeners.

It is considered below (a) why all the language groups showed neutral responses in the two-syllable block, and (b) why the overall responses of all the groups were in the direction of an initial-stress bias in the three-syllable block though the Japanese listeners' bias was the weakest.

<Two-Syllable Natural>

The correct response rates of all the language groups varied between 60%

and 70% for both of the stress patterns, which was above chance level. One possible interpretation of the result is that all the three language groups were able to hear across-stimuli pitch differences. As shown in Table 6 (Section 4.2), the mean F0 difference between Vowel 1 (V1) and Vowel 2 (V2) was 7 Hz in the -*ing* forms, e.g., *DOmi*-, and that between V1 and V2 was 4 Hz in the -*ion* forms, e.g. *domi*-. That is, their slopes were similar. Nonetheless, they were different in their pitch ranges: the F0 contour of *DOmi*- was in the pitch range about 10 Hz higher than that of *domi*-. That is, 'across-stimuli' evidence based on pitch range was available for participants.

Across-stimuli durational evidence was also available. As already shown in Table 5, Section 4.2, the initial syllable (Syllable 1) of the -*ing* form (*DOmi*-) was about 30 ms longer than that of the -*ion* form (*domi*-). They were probably able to hear the difference at least in this condition.

<Three-Syllable Natural>

By adding the third syllable to the stimuli, a bias towards the -*ing* forms (*DOminating*) suddenly emerged, especially in the responses of the native English listeners and Seoul Korean listeners. One possible account is that the addition of the third syllable allowed the stimuli extracted from the -*ing* forms to provide within-stimulus evidence based on pitch. As shown in Table 6, Section 4.2., the -*ing* stimuli had clear pitch difference of approximately 20 Hz between V1 and V3 and approximately 15 Hz between V2 and V3. In contrast, the F0 values of the three vowels in the -*ion* stimuli stayed almost constant. That is why the -*ing* stimuli obtained more correct responses than the -*ion* stimuli.

<Three-Syllable Slight Slanting>

What is noticeable in the slightly slanting block is a drastic drop by 25% to 30% in the correct response rates of the -*ion* forms in all language groups, and the drop was especially noticeable in the English listeners' and the Seoul Korean listeners' responses. It is only the Japanese listeners' correct response rates that barely reached the chance level. Another thing to note is that although the three language groups showed relatively high correct response rates of the -*ing* forms, the Japanese listeners' are the lowest. The question is whether those are all explained by only the acoustic characteristics of the stimuli or some other explanations are necessary.

In this block, all stimuli had the same flat F0 contour regardless of their original stress patterns (see Figures 10 and 11 in Section 4.2). The left edge of the contour was set at 90 Hz and the right edge was set at 85 Hz. The F0 value averaged across the V1 period was slightly lower than 90 Hz and that averaged across the V3 period was slightly higher than 85 Hz, and the difference between the two vowel periods was actually smaller than 3 Hz. This pitch condition provided neither across-stimuli nor within-stimulus evidence.

Although pitch-related evidence was scarce, the -*ing* stimuli had within-stimulus durational evidence. As seen in Table 5, the mean duration of the first foot, i.e. the mean duration of Syllable 1 (S1) and that of Syllable 2 (S2) combined together, was 75 ms longer than the duration of Syllable (S3) (S1 + S2 = 280 ms; S3 = 205 ms). The correct response rates of the three language groups were relatively high for the -*ing* stimuli probably because this within-stimulus durational evidence was available. Nonetheless, the correct rates of the Japanese listeners were slightly lower than those of the other two language groups. It is because the Japanese listeners were influenced by the predominance of antepenultimate-mora accent in their L1 and were biased towards the -*ion* forms but not towards the -*ing* forms.

In the -*ion* stimuli, the difference between the first foot and S3 was only 26 ms (S1 + S2 = 249 ms; S3 = 223 ms), which was not probably sufficient for the listeners to count as within-stimulus durational evidence. As a result, the correct response rates of the -*ion* forms were low regardless of the language difference. Nonetheless, the Japanese listeners' correct rates were much higher than the other two groups'. This is again because the Japanese listeners were influenced by the predominance of antepenultimate-mora accent in their L1.

Another thing to note regarding durational cues is that the participants were not able to pay attention to the across-stimuli durational difference in this block. As already mentioned above, S1 in the -*ing* forms was 30 ms longer than S1 in the -*ion* forms (see Table 5). S3 also had a difference of about 20 ms between them (see Table 5). If the listeners had been able to pay attention to this across-stimuli durational evidence, they could have given more correct responses to the -*ion* forms. Because this block had stimuli consisting of three syllables, they had to process quite a few pieces of information all at once and they probably could not afford to pay attention to the across-stimuli durational differences.

Although I listed several factors in Table 2 of Section 2.1.3 as the frequency-related knowledge of English listeners that are possibly affecting their responses, one of them turned out not to be contributing to them at all: the knowledge in (iii) that the frequency of the -*ing* forms tends to be lower than that of the -*ion* forms was already proven not to have affected their responses by the ANOVA shown in the previous section.

It is not clear from this task whether the English listeners were influenced by the predominance of initial stress in their vocabulary, because their response patterns have been accounted for even without taking it into consideration, though the result that they preferred initial stress complies with the frequency factor.

5. Summary and conclusion

The results of two tasks with truncated stimuli were reported in this article to test the hypothesis that Japanese listeners are affected by the predominance of antepenultimate-mora accent in their native language even when they judge the stress patterns of English words. One of the tasks used a set of stimuli called Group 1 consisting of only the initial syllables of trochaic nouns, e.g., *TRANS*- from *TRANSplant*, and those of iambic verbs, .e.g., *trans*- from *transPLANT*. The other employed a set of stimuli called Group 2 consisting of the initial two or three syllables of four-syllable suffixed words such as *DOminating* and *domiNAtion*. In those tasks, the participants were asked to identify the words from which the stimuli were extracted. The task with the Group 1 stimuli did not obtain the result predicted by the hypothesis while the task with the Group 2 stimuli in the three-syllable conditions obtained a result in line with the hypothesis. In those conditions, the Japanese listeners showed the least bias towards initial stress (-*ing* forms) among all the three language groups.

Although it has been argued throughout this article that Japanese listeners' perception of English stress is influenced by the antepenultimate accent rule (AAR) in their L1, let me also add that there is a possibility that what matters is in fact the Latin accent rule (LAR) which is also known to operate in the accent assignment of Japanese words (Kubozono, 1996, 2002, 2006, 2015; Shinohara, 2000). LAR is a rule based on syllable weight, which states that the penultimate syllable is stressed if it is heavy, otherwise the antepenultimate syllable is stressed, and it is operative not only in Romance languages but also in English and other languages (Hayes, 1995; Mester, 1994). Kubozono (1996, 2002, 2006, 2015) notices similarity between the predictions made by AAR and those by LAR for Japanese loanword accent assignment: the two rules make the same predictions for most of the cases, and the only exception is when words end with a sequence of a light and a heavy syllable. Examples are *ka⌉a.di.gan* predicted by LAR and *kaa.di⌉.gan* predicted by AAR. The accent patterns predicted by AAR and those predicted by LAR actually coexist in Japanese, and both *ka⌉a.di.gan* and *kaa.di⌉.gan* are the possible accent patterns for the Japanese loanword originating from the English word *cardigan*. Kubozono (2002) further reports that the pattern predicted by LAR in fact outnumbers the pattern predicted by AAR, and Japanese is now changing towards the direction of LAR (Kubozono, 2002, 2015). Given this, consider one of the stimuli with the *-ion* suffix used in this study: *dominátion*. The accent location of the Japanese loanword counterpart of this form *do. mi. ne⌉e. ʃon* is the syllable that contains the antepenultimate mora at the same time that it is the penultimate syllable that is heavy. Therefore, the hypothesis based on LAR can also make a correct prediction. It has been also reported by Ishikawa (2007) that when asked to predict a stress position in pseudo English verbs that end with a heavy syllable (CVCC) with different inflectional suffixes, e.g., *formand*, *formands*, *formanded*, *formanding*, Japanese learners of English are more likely to put stress on the stem-final syllables when verbs were with a syllabic suffix, e.g., *formá̲nded* and *formá̲nding*, and they were more likely to place stress on the initial syllables when they were with a consonantal suffix or without a suffix, e.g., *fó̲rmand*, *fó̲rmands*. Sugahara

(2016b) also obtained a similar result in her questionnaire study with real words. The initial syllables in Ishikawa's (2007) pseudo verbs with a consonantal suffix and those without a suffix e.g., *fórmand*, *fórmands*, and the stem-final syllables in the forms with syllabic inflectional suffixes, e.g., *formánded* and *formánding*, are all heavy penultimate syllables, which are predicted by LAR. In English, however, LAR is the rule that is in principle applied to nouns, and it is natural to consider the Japanese speakers' overapplication of LAR to English verbs to have its origin in Japanese. A further investigation is necessary regarding this point.

Acknowledgement

This work was supported by JSPS KAKENHI (Grant-in-Aid for Scientific Rsearch (C)), Grant Number 24520558, which was awarded to the author, and by NINJAL collaborative research project 'Cross-linguistic Studies of Japanese Prosody and Grammar'. An earlier version of this study was presented at *Phonology Festa 11* held at Ritsumeikan University in March, 2016.

Notes

1 In English, only syllables with primary stress may bear a nuclear pitch accent. Although the tonal shapes of a nuclear pitch accent may take various forms, about 95% of all the occurrences of pitch accents in English are those with an H accent tone (Dainora, 2006).

2 The number of such minimal pairs in English is small.

3 They do not include trisyllabic verbs ending with affixes that bear secondary stress such as -*àte*, e.g., *dóminàte*, and -*ìze*, e.g., *prívatìze* because the final syllables of trisyllabic words with antepenultimate stress counted by Hammond were all unstressed.

4 When penultimate, they tend to be heavy.

5 Final syllables with stress tend to be superheavy (trimoraic), closed by a consonant or by a consonant cluster, unless they contain a long vowel, e.g., *gùarantée*.

6 Zamma (2007) gives a summary of Class I suffixes referring to their stress patterns.

7 I do not take into the factor of neighborhood density when making predictions here. It is because the competitors evoked in the minds of listeners soon after they hear the truncated stimuli are always 'two', i.e, the two options visually presented to them as possible identities of the stimuli. This is different from ordinary environments where all or a subset of the competitors sharing the same initial syllable in the lexicon are evoked in listener's minds.

8 When English words are introduced into Japanese as loanwords, an epenthetic vowel is inserted after consonants that are not immediately followed by a vowel because Japanese phonotactics in principle allows only CV syllables. The only exceptions are nasal coda CVN and the initial half of a geminate consonant CVQ.

9 At the intonational phrase-final position, the last LH tones is replaced with a falling tone in a declarative context.

10 All the inflected forms of a single word, e.g., *transplant*, *transplants*, *transplanted*, *transplanting*, belong to a single 'lemma'. Therefore, the lemma frequency of *transplant* (verb), for example, is the sum frequency of all the inflected forms.

11 The frequency information stored in CELEX is the frequency count of word occurrences in the COBUILD corpus of the University of Birmingham based on various types of written sources and some speech sources.

12 In Sugahara (2016a), I reported that all of them were exchange students studying at Doshisha for a year or half a year, which was an error. Their affiliation, however, does not affect the results and the analyses presented in Sugahara (2016a).

13 In Sugahara (2016a), I reported that all but one were from the US, which was also an error. All the varieties of English presented here (American English, British English spoken in England, Australian English, Canadian English) share the same stress patterns for the noun-verb pairs used in both Sugahara (2016a) and the current study. Therefore, the variety differences among the native English speakers are not expected to affect the results and the analyses.

14 The formula used to calculate the '*c*' value is the following: $c = -(z(\text{Hit})+z(\text{FA}))/2$, where z(Hit) is the z value of a 'Hit' rate, i.e., the rate of correct responses to the stimuli originally produced as iambic, e.g., *trans*(*PLANT*), and z(FA) is the z value of an 'False Alarm' rate, i.e., the rate of incorrect (iambic) responses to the stimuli originally produced as trochaic. When those rates are 1 or 0, z scores cannot be

calculated. In such cases, they were converted by using the following calculus proposed by Macmillan & Creelman (2005): $z(1)$ was converted to $z(1\text{-}0.5/n)$ and $z(0)$ was to $z(0.5/n)$, where n stands for the number of the relevant iambic-trochaic pairs (see also Kawahara & Shinohara, 2015). Since there were only five pairs in the Group 1 stimuli, $z(1)$ was converted to 0.9 and $z(0)$ was converted to 0.1.

15 The Japanese listeners' bias towards final stress, however, was significantly stronger than the Seoul Korean listeners' one (Sugahara, 2016a).

16 The calculus to obtain the *c* values is shown in footnote 14.

References

Baayen, R. H. & Piepenbrock, R. & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.

Beckman, M. E. (1986). *Stress and non-stress accent*. Dordrecht: Foris Publications.

Boersma, P. & Weenink, D. (2013). *Praat: doing phonetics by computer* [Computer program]. Version 5.3.51, retrieved 2 June 2013 from http://www.praat.org/.

Clopper, C. G. (2002). Frequency of stress patterns in English: a computational analysis. *IULC Working Papers, 2*.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: evidence from native and non-native listeners. *Language and Speech, 45*, 207-228.

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

Dainora, A. (2006). Modeling intonation in English: a probabilistic approach to phonological competence. In L. Goldstein, D. H. Whalen & C. T. Best (Eds.), *Laboratory Phonology 8* (pp. 107-132). Berlin, New York: Mouton de Gruyter.

Guion, S. G. (2005). Knowledge of English word stress patterns in early and late Korean-English bilinguals. *Studies in Second Language Acquisition, 27*, 503-533.

Guion, S. G., Harada, T., and Clark, J. J. (2004). Early and late Spanish-English bilinguals' acquisition of English word stress patterns. *Bilingualism: Language and Cognition, 7*, 207-226.

Hammond, M. (1999). *The phonology of English*. Cambridge: Cambridge University Press.

Hayes, B. (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.

Ishikawa, K. (2007). Grammatical class and rhythmic context: English stress assignment by Japanese students. *JACET Journal, 44*, 29-42.

Ishikawa K, & Nomura, J. (2008). Word-stress placement by native speakers and Japanese learners of English. *Proceedings of Interspeech 2008* (pp .1955-1958). Brisbane.

Jun, S-A. (1996). *The phonetics and phonology of Korean prosody*. New York, London: Garland Publishing.

Jun, S-A. (1998). The Accentual Phrase in the Korean prosodic hierarchy. *Phonology, 15*, 189-226.

Jun, S-A. (2005). Korean intonational phonology and prosodic transcription. In S-A. Jun (Ed.), *Prosodic typology* (pp. 201-229). Oxford: Oxford University Press.

Jun, S-A. (2006). Intonational phonology of Seoul Korean revisited. In T. Vance & K. Jones (Eds.), *Japanese Korean Linguistics 14* (pp.15-26). Stanford: CSLI.

Kawahara, S., & Shinohara, K. (2015). Iconic inferences about personality: from sounds and shapes. In M. K. Hiraga, W. J. Herlofsky, K. Shinohara & K. Akita (Eds.), *Iconicity: East meets West* (pp.57-70). Amsterdam: John Benjamins Publishing Company.

Kubozono, H. (1996). Syllable and accent in Japanese: evidence from loanword accentuation. *The Bulletin* (The Phonetic Society of Japan) *211*, 71–82.

Kubozono, H. (2002). Prosodic structure of loanwords in Japanese: syllable structure, accent and morphology. *The Journal of the Phonetic Society of Japan, 6 (1)*, 79–97.

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

Kubozono, H. (2015). Loanword phonology. In H. Kubozono (Ed.), *The handbook of Japanese phonetics and phonology*. Berlin: De Gruyter Mouton.

Macmillan, N. A., & Creelman, C. D. (1990). Response bias: characteristics of detection theory, threshold theory and "nonparametric" indexes. *Psychological Bulletin, 107(3)*, 401-413.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: a user's guide* (2nd ed.). Mahwah: Lawrence Erlbaum Associates Publishers.

McCawley, J. D. (1968). *The phonological component of a grammar of Japanese*. The Hague: Mouton.

Mester, A. (1994). The quantitative trochee in Latin. *Natural Language and*

Linguistics Theory, 12, 1-62.

Ou, S.-C. (2010). Taiwanese EFL learners' perception of English word stress. *Concentric: Studies in Linguistics, 36*, 1-23.

Ou, S.-C. & Guo, Z-C. (2015). The effect of stress on English word recognition by native speakers of typologically different languages. *Proceedings of ICPhS 2015*.

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

Schaefer, V., & Darcy, I. (2014). Lexical function of pitch in the first language shapes cross-linguistic perception of Thai tones. *Laboratory Phonology, 5*, 489-522.

Shinohara, S. (2000). Default accentuation and foot structure in Japanese: evidence from Japanese adaptations of French words. *Journal of East Asian Linguistics, 9*, 55-96.

Sugahara, M. (2011). Identification of English primary stress and bias toward strong word-initial syllables: native vs. Japanese listeners. *Proceedings of ICPhS 2011* (pp. 1918-1921).

Sugahara, M. (2016a). Is Japanese listeners' perception of English stress influenced by the antepenultimate accent in Japanese? Comparison with English and Korean listeners. *Doshisha Studies in English*, *96*, 61-111.

Sugahara, M. (2016b). The influence of L1 lexical prominence systems on the judgement of stress locations in written English words: comparison between Japanese and Seoul Korean speakers. Poster presentation at the *24th Japanese Korean Conference*.

Sugito, M. (1995). *Osaka/Tokyo akusento onsei jiten* [Pronunciation dictionary of Osaka/Tokyo accent]. Tokyo: Maruzen.

Tanaka, S-I. (2009). *Osaka hogen gairaigo-no akusento-to shiki-ni tuite* [On accent and shiki in Osaka Japanese loanwords]. *Proceedings of the 138th meeting of Linguistic Society of Japan* (pp.214-219).

Van Leyden, K., & van Heuven, V. J. (1996). Lexical stress and spoken word recognition: Dutch versus English. In M. den Dikken & C. Cremers (Eds.), *Linguistics in the Netherlands* (pp. 159-170). Amsterdam: John Benjamins.

Zamma, H. (2007). Categorical and non-categorical variation in English stress assignment. *Phonological Studies*, *10*, 49-56.

Appendix 1. Duration (ms): Group 2

		Vowel duration (ms)			Syllable duration (ms)		
		Vowel1	Vowel2	Vowel3	Syllable1	Syllable2	Syllable3
activa	ting	112	58	125	180	131	172
	tion	87	49	146	155	113	220
agita	ting	122	59	141	122	133	238
	tion	89	51	130	89	130	265
alloca	ting	117	75	105	117	111	213
	tion	90	78	127	90	117	238
calcula	ting	151	58	112	171	166	162
	tion	119	52	126	137	147	194
captiva	ting	137	36	129	221	101	174
	tion	121	44	122	199	111	174
celebra	ting	87	56	126	268	96	219
	tion	66	37	151	224	76	244
complica	ting	79	51	121	181	114	208
	tion	58	68	131	172	112	230
concentra	ting	73	70	131	200	145	229
	tion	81	78	113	180	147	237
conjuga	ting	84	39	126	236	92	172
	tion	75	39	152	207	83	206
dedica	ting	89	51	139	99	83	247
	tion	70	48	124	81	74	270
domina	ting	117	49	122	127	108	173
	tion	102	32	131	110	89	190
educa	ting	89	58	132	89	131	234
	tion	68	52	123	68	132	236
estima	ting	85	24	116	189	82	185
	tion	81	28	112	174	94	189
indica	ting	87	40	136	155	63	247
	tion	65	34	121	123	75	239
medica	ting	74	58	127	194	91	227
	tion	58	40	134	135	77	254
mitiga	ting	68	76	132	141	103	205
	tion	72	63	158	113	100	227
motiva	ting	116	62	139	241	94	199
	tion	95	54	118	202	83	178
naviga	ting	130	56	143	226	77	225
	tion	83	53	152	172	92	237
propaga	ting	82	57	139	145	111	201
	tion	62	45	130	151	102	205
prosecu	ting	100	37	93	187	126	209
	tion	99	47	125	145	126	242
termina	ting	71	49	121	153	108	164
	tion	58	27	142	135	84	206

Appendix 2. F0 and intensity (averaged across each vowel period): Group 2

		Vowel 1		Vowel 2		Vowel 3	
		F0 (Hz)	Intensity (dB)	F0 (Hz)	Intensity (dB)	F0 (Hz)	Intensity (dB)
activa	ting	105	61	87	60	79	60
	tion	94	60	87	59	92	62
agita	ting	100	63	96	59	82	56
	tion	92	62	85	55	96	60
alloca	ting	96	62	93	62	76	57
	tion	90	61	87	60	88	59
calcula	ting	93	60	80	58	76	56
	tion	86	60	87	57	84	57
captiva	ting	96	60	87	59	77	59
	tion	92	60	88	57	90	60
celebra	ting	103	64	95	60	76	59
	tion	93	62	89	59	89	62
complica	ting	99	67	92	57	79	56
	tion	102	67	95	57	94	61
concentra	ting	103	66	90	65	78	57
	tion	92	63	84	63	86	59
conjuga	ting	99	65	87	56	79	58
	tion	90	64	87	55	87	60
dedica	ting	105	63	99	59	82	57
	tion	92	59	83	57	90	60
domina	ting	102	67	96	68	81	64
	tion	91	65	88	66	90	65
educa	ting	99	61	96	58	82	57
	tion	95	59	93	59	93	60
estima	ting	99	61	88	62	75	62
	tion	97	61	94	62	89	66
indica	ting	104	68	98	61	82	57
	tion	100	68	85	57	95	59
medica	ting	98	63	93	59	82	57
	tion	89	63	95	56	88	59
mitiga	ting	93	64	88	58	75	54
	tion	85	62	88	58	86	59
motiva	ting	102	65	98	62	80	59
	tion	88	60	87	59	90	61
naviga	ting	100	65	93	63	75	55
	tion	93	66	92	62	94	63
propaga	ting	99	63	93	59	79	55
	tion	86	60	88	57	85	57
prosecu	ting	103	64	97	57	84	56
	tion	92	60	91	56	92	60
termina	ting	108	68	92	67	78	62
	tion	91	65	88	65	89	64

Appendix 3. First and second formants (F1 and F2), obtained from the mid point of each vowel period: Group 2

		Vowel 1		Vowel 2		Vowel 3	
		F1	F2	F1	F2	F1	F2
activa	ting	760	1682	427	1674	439	2017
	tion	746	1646	432	1614	452	2056
agita	ting	816	1613	363	1926	439	2097
	tion	677	1647	367	1799	422	2058
alloca	ting	767	1413	540	1063	478	2006
	tion	770	1399	490	1071	420	2076
calcula	ting	816	1372	523	1544	512	1798
	tion	687	1259	488	1615	450	1857
captiva	ting	777	1550	436	1570	485	1922
	tion	763	1497	434	1456	444	1956
celebra	ting	625	1468	554	1097	512	1672
	tion	607	1372	560	1026	505	1622
complica	ting	695	1083	461	1149	459	2017
	tion	702	1114	411	923	453	1953
concentra	ting	665	1160	332	717	501	1640
	tion	685	1256	562	2465	486	1737
conjuga	ting	661	1200	317	2164	452	1989
	tion	673	1175	347	2111	434	1979
dedica	ting	520	1762	445	1870	458	2040
	tion	506	1731	447	1903	437	2078
domina	ting	689	1207	534	1488	522	2201
	tion	712	1237	533	1609	558	2286
educa	ting	528	1909	366	2137	455	2000
	tion	509	1876	323	2101	431	1998
estima	ting	662	1588	438	1483	487	2003
	tion	654	1663	447	1528	526	2101
indica	ting	528	2751	379	2086	451	2075
	tion	576	2852	397	2099	426	2088
medica	ting	583	1765	443	1994	434	2069
	tion	544	1690	429	1941	426	2124
mitiga	ting	455	1848	401	2085	465	2070
	tion	488	1883	369	2082	444	2035
motiva	ting	525	1305	416	1526	477	1905
	tion	540	1128	430	1468	451	1930
naviga	ting	689	1917	485	1729	451	2071
	tion	670	1576	503	1796	450	2095
propaga	ting	720	1136	492	1598	463	2053
	tion	701	1151	491	1635	442	1989
prosecu	ting	685	1142	394	1741	380	2031
	tion	664	1160	424	1691	357	1930
termina	ting	679	1628	544	1735	544	2081
	tion	691	1498	546	1613	501	2241