論 文

On Japanese EFL Learners' Word Stress Assignment in Derived Words

Takahiro Ioroi

1. Introduction

Placement of word stress in polysyllabic words is an area of English pronunciation with which Japanese EFL learners often show difficulty. Imbalanced combinations of primary stress and secondary stress result in the absence of rhythm in their spoken output. The present study examines a set of words uttered by 22 Japanese EFL learners in comparison with data based on pronunciation models from a commonly available online dictionary. Specifically, the discussion centers on how primary stress and secondary stress are realized in the learners' interlanguage and on whether morphology has any effects on the manifestation of stress in complex words. To this end, a method that represents and compares acoustic measurements is proposed.

2. Literature review

2.1 Lexical stress and its phonetic correlates

Given a word that consists of more than a single syllable, categorical terms such as 'primary stressed', 'secondary stressed', and 'unstressed' are employed in phonology to indicate distinctions in the loudness of the syllables in the word. One of the greatest discoveries that has been achieved in phonological theory over the several decades is that locations of word stress are mostly predictable: namely that they are rule-governed and reflect the speaker's implicit/procedural knowledge.

However, what is referred to as linguistic stress is not in fact a single, uniform entity. As Laver (1994: 511) points out, stress is a gradient phenomenon in phonetics and "the phonetic realization of any syllable can be said to show a greater or less degree of stress relative to the manifestation of some other syllable." As for acoustic correlates of stress, it is generally accepted that there are certain elements that give rise to differences in stress levels. Mattys (2000: 254) states that "it is suprasegmental variables, such as frequency, duration, and amplitude, that determine stress perception." Laver (1994: 511) suggests four dimensions: quality, duration, loudness and pitch. Galves et al. (2002) propose an approach to rhythmic classes based on sonority. Ramus et al. (1999) used the duration of vocalic intervals and consonantal intervals in an attempt to investigate rhythm perception by infants with different linguistic backgrounds.[1]

2. 2 The influence of morphology on pronunciation

Intensive research on the influence of morphology, i.e., word structure, on the placement of lexical stress appeared in Chomsky and Halle (1968) and then evolved into the theory of Lexical Phonology in the 1970s through the late 1980s (Mohanan 1982, Kiparsky 1982). A series of studies within that framework uncovered complex interactions between morphology and phonology that were assumed to take place in the lexical component of grammar. One of the important claims of Lexical Phonology was that there are two classes of affixes: Class I and Class II. Class I includes such suffixes as *-ation, -ity, -ize,* and *-ic* while Class II includes *-ment, -ness, -ing,* and *-able,* to name a few. Typically, Class I suffixes affect the location of word stress. In contrast, Class II suffixes trigger no such change after suffixation takes place. L1 English speakers distinguish these classes of suffixes and assumedly manipulate rules of stress assignment and relevant morpho-phonological alternations. L2 English learners need to acquire the implicit knowledge of this component as part of vocabulary knowledge. However, there seems to be little research into what operates on their interlanguage stress assignment.

3. Data collection

3. 1 Subjects

Twenty-two female undergraduate EFL students participated this experiment.

3. 2 Stimuli

The stimuli were eight nouns composed of four syllables: *admiration, declaration, explanation, invitation, disappointment, establishment, encouragement,* and *entertainment. Admiration, declaration, explanation* and *invitation* are formed from *admire, declare* and *explain* by means of the noun-forming suffix *-ation* while *disappointment, establishment, encouragement* and *entertainment* are derived by the noun-forming suffix *-ment,* with *disappoint, establish, encourage* and *entertain* as their bases. Table 1 summarizes the makeup of the two sets of complex words.

Table 1. Stimuli sorted by type of suffixation

Class I suffixation (Stress-shifting)	Class II suffixation (Stress-neutral)
àdmirátion < admíre	dìsappóintment < dìsappóint
dèclarátion < decláre	encóuragement < encóurage
èxplanátion < expláin	èntertáinment < èntertáin
ìnvitátion < invíte	estáblishment < estáblish

Table 2 shows that six of the eight items carry a secondary stressed syllable and the remaining two exhibit primary stress only:

Table 2. Presence or absence of secondary stress

Words with secondary stress	Words with primary stress only
àdmirátion, dèclarátion, èxplanátion, ìnvitátion, dìsappóintment, èntertáinment	encóuragement, estáblishment

3. 3 Procedure

The recording of the 22 subjects reading the eight words aloud were conducted in a language laboratory using a Sanyo ICR-S290RM recorder with an external microphone at the sampling rate of 16 kHz. The recorded utterances were saved in the WAV format.

3. 4 Reference speech sounds

As a source of reference speech sounds, the online edition of the *Sanseido Exceed English-Japanese Dictionary* (*Exceed* hereafter) was employed.[2] This dictionary served as a suitable source because the pronunciation models for the selected eight words are based on female speakers and the sound files are already in the WAV format and can be directly read into the speech analysis software.

3. 5 Measurement

For each word, the amplitude (dB), pitch (Hz), and duration (ms) on the primary-stressed, secondary-stressed, and unstressed syllables were measured on WaveSurfer.[3]

4. Findings and discussion

4. 1 An overview

There were eight tokens in the set sampled from *Exceed* (8x1=8). The sample set from the 22 learners had 176 tokens (8x22=176). For each of the tokens, the ratio between the primary-stressed syllable and the unstressed syllable (PS:NS) and the ratio between the secondary-stressed syllable and the unstressed syllable (SS:NS) were calculated with respect to the three dimensions of word stress introduced in the previous section. Tables 3 and 4 present the PS:NS and SS:NS ratios from *Exceed* and the Japanese EFL learners for the eight lexical items with respect to the three dimensions introduced above. Note that the cells for the learners' SS:NS ratios are shaded in gray, as two of the eight items carry primary stress only. Figures 1 and 2 visualize these results as sequential line charts.

Table 3. PS:NS and SS:NS ratios for amplitude, pitch, and duration (*Exceed*)

Items	PS:NS (amplitude)	SS:NS (amplitude)	PS:NS (pitch)	SS:NS (pitch)	PS:NS (duration)	SS:NS (duration)
admiration	1.06	1.08	1.10	1.83	1.97	1.25
declaration	0.94	0.94	1.13	2.22	2.00	1.23
explanation	1.10	1.07	1.04	2.88	2.01	0.79
invitation	1.06	1.10	1.13	5.90	2.88	2.19
disappointment	1.05	1.02	1.01	7.34	4.59	1.83
entertainment	1.11	1.03	1.07	4.16	2.60	1.66
encouragement	1.01		2.51		0.99	
establishment	1.09		1.34		1.60	
Mean	1.05	1.04	1.29	4.06	2.33	1.49
Mean (*-ation*)	1.04	1.05	1.10	3.21	2.22	1.37
Mean (*-ment*)	1.07	1.03	1.48	5.75	2.45	1.75

Table 4. PS:NS and SS:NS ratios for amplitude, pitch, and duration (Japanese EFL learners)

Items	PS:NS (amplitude)	SS:NS (amplitude)	PS:NS (pitch)	SS:NS (pitch)	PS:NS (duration)	SS:NS (duration)
admiration	1.14	1.04	1.10	2.20	2.07	0.79
declaration	1.11	0.98	1.06	2.47	2.18	0.86
explanation	1.01	0.92	1.10	1.60	1.43	0.43
invitation	1.13	1.02	1.11	4.75	3.42	1.80
disappointment	1.09	0.95	1.22	4.66	4.23	1.24
entertainment	1.07	0.99	1.10	2.85	2.48	1.21
encouragement	1.12		1.20		1.73	
establishment	1.10		1.34		1.74	
Mean (all)	1.10	0.98	1.15	3.09	2.41	1.06
Mean (*-ation*)	1.10	0.99	1.09	2.76	2.28	0.97
Mean (*-ment*)	1.10	0.97	1.22	3.76	2.55	1.23

Figure 1. The PS:NS and SS:NS ratios for amplitude, pitch, and duration (*Exceed*)

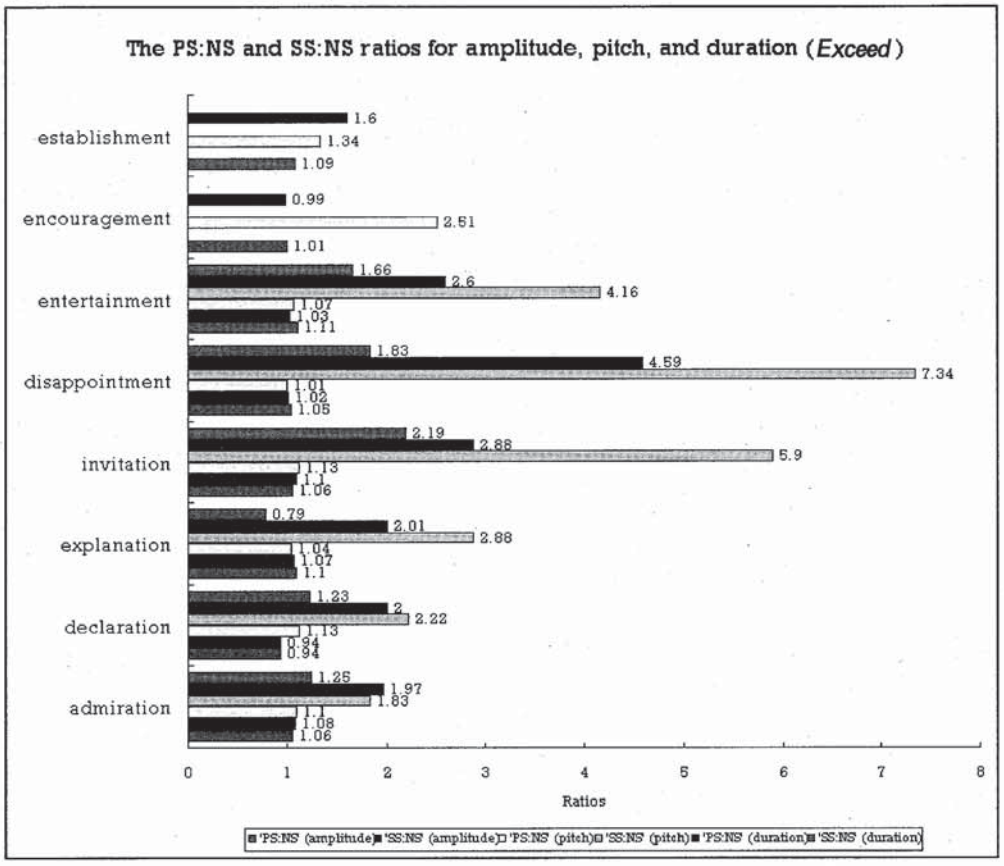
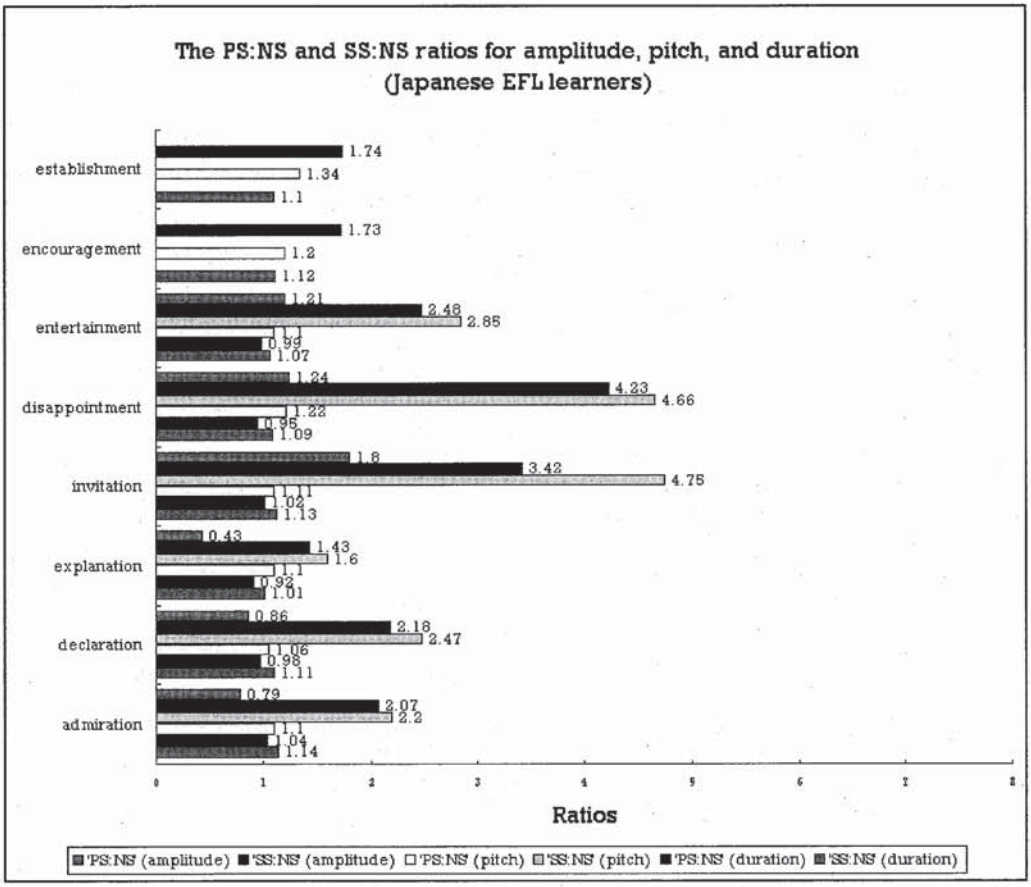


Figure 2. The PS:NS and SS:NS ratios for amplitude, pitch, and duration (Japanese EFL learners)



The results above address two important features that differentiate primary stressed syllables and secondary stressed syllables. One is the PS:NS durational ratios and the other is the SS:NS ratios in pitch. Note that every item reveals a distinctively higher PS:NS ratio regarding duration than the other two dimensions, indicating that syllables are significantly lengthened when they receive primary stress. On the contrary, SS:NS ratios exhibit far higher scores in terms of pitch. This means that secondary stressed syllables attract higher pitch than primary stressed syllables do.

Regarding the PS:NS ratios, both of the data from *Exceed* and that from the learners show scores greater than 1.0, conforming to the natural expectation that the syllables on which primary stress fall is stronger than the unstressed syllables. It should be added that the data from the learners presents a score greater than that from *Exceed* (E: 1.05, J: 1.10). With respect to pitch, the data shows an average score of 1.29, which is greater than that from the learners (E: 1.29, J: 1.15). In contrast, *Exceed* exhibits a smaller score than the data from the learners does with respect to duration (E: 2.33, J: 2.41).

As for the SS:NS ratios, the data from the subjects reveals lower scores than *Exceed* does in all of the three dimensions: amplitude (E: 1.04, J: 0.98), pitch (E: 4.06, J: 3.09), duration (E: 1.49, J: 1.06). These results help to confirm the impression that Japanese EFL learners, whose native language is categorized as a syllable-timed language (Abercrombie 1967), are markedly less proficient in producing secondary stress strong enough to conform to English rhythmic patterns. Table 5 displays the data sorted according to the distinction between the two derivational suffixes.

Table 5. Differences in the PS:NS/SS:NS ratios according to suffix classes

Exceed	PS:NS (amplitude)	SS:NS (amplitude)	PS:NS (pitch)	SS:NS (pitch)	PS:NS (duration)	SS:NS (duration)
Mean (*-ation*)	1.04	1.05	1.10	3.21	2.22	1.37
Mean (*-ment*)	1.07	1.03	1.48	5.75	2.45	1.75
Difference	▲0.03	0.02	▲0.38	▲2.54	▲0.23	▲0.38
Japanese EFL learners	PS:NS (amplitude)	SS:NS (amplitude)	PS:NS (pitch)	SS:NS (pitch)	PS:NS (duration)	SS:NS (duration)
Mean (*-ation*)	1.1	0.99	1.09	2.76	2.28	0.97
Mean (*-ment*)	1.1	0.97	1.22	3.76	2.55	1.23
Difference	0.00	0.02	▲0.13	▲1.00	▲0.27	▲0.26

Neither *Exceed* nor the Japanese EFL learners show much difference between the two sets of words with respect to amplitude. In view of duration, on the other hand, the suffix *-ment* shows its predominant tendency over the suffix *-ation*. In addition, differences between the two

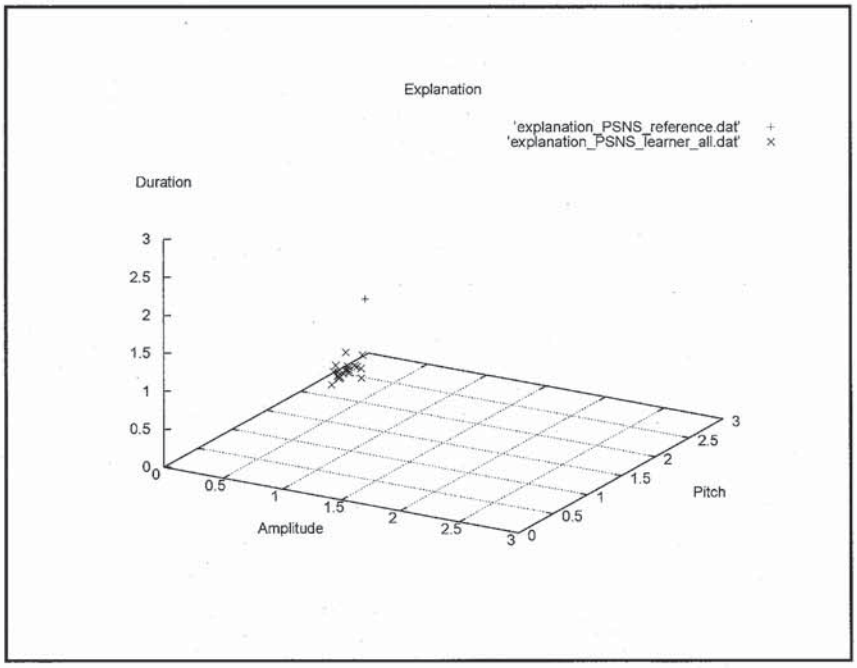
sets of morphologically complex words are greater in the data from Exceed than in the data from the Japanese EFL learners across the dimensions except for the durational PS:NS ratios. At present it is not clear whether these differences are caused by something ascribable to the two suffix classes or due to unpredictable, idiosyncratic properties of the individual words in question.

4. 2 The locations of individual tokens in the prosodic space and their distances from the reference points

Suppose that we have a three dimensional space with its three axes representing amplitude, pitch, and duration, respectively. Let us call it the 'prosodic space'. Given the noun *explanation,* for example, the PS:NS ratios obtained from *Exceed* and the 22 learners can be displayed in the prosodic space, as in Figure 3. The symbol + indicates the location of the sample from *Exceed* and the ×'s stand for the 22 tokens from the learners.

Figure 3 represents the location of each of the samples from both sources.

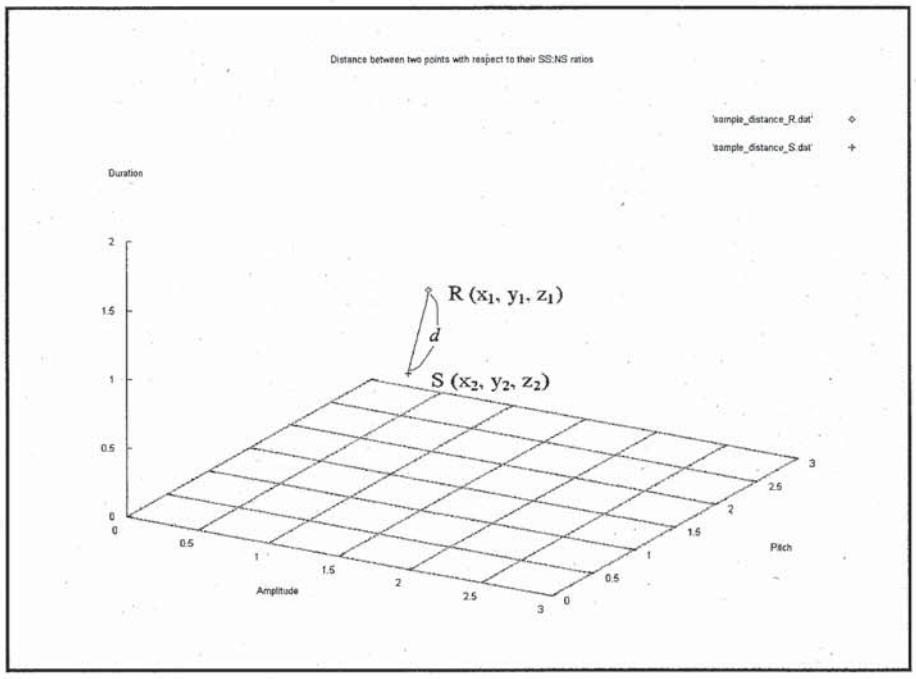
Figure 3. Locations of PS:NS ratios in the prosodic space in regard to *explanation*



It is without doubt that most of the data from the learners are concentrated on a particular area and look like a single cluster while the reference data holds an optically visible distance from the area crowded with ×'s.

Let us introduce the notion 'prosodic' distance. Given a pair of points R (x_1, y_1, z_1) and S (x_2, y_2, z_2) obtained through the above-mentioned procedure, we can locate them in the prosodic space as illustrated in Figure 4.

Figure 4. The distance *d* between R and S in the prosodic space



The distance between the reference point and the given token can be expressed as the geometrical distance *d* indicated by the straight line between R and S. The value of *d* is obtained by the following expression:

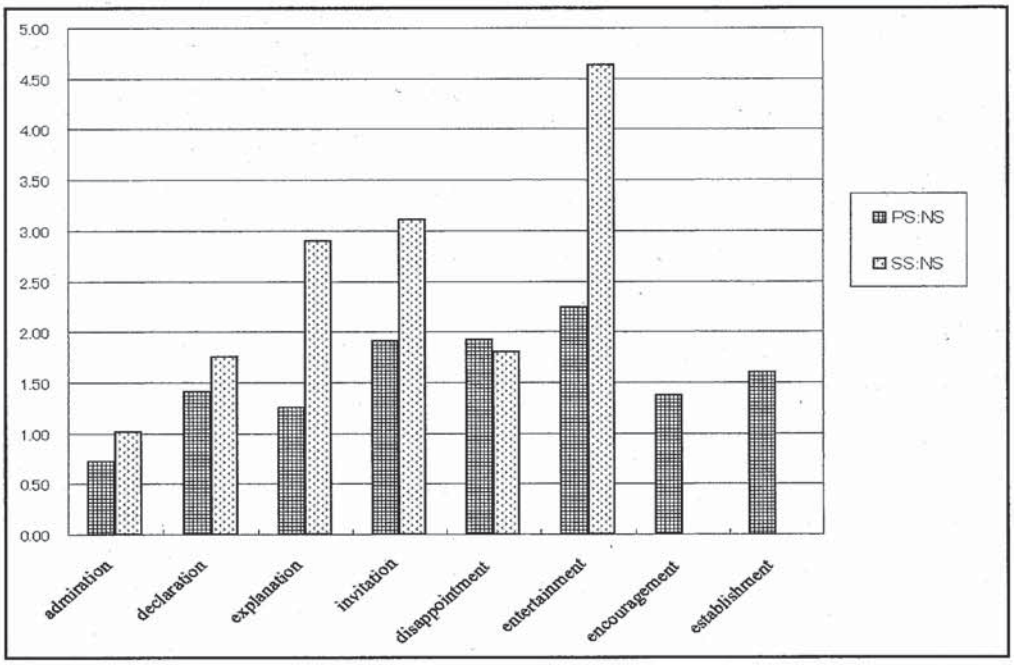
$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

Naturally, as the value of *d* increases, the distance from the reference point becomes greater. This formula produced the data in Table 6. They are the mean scores of the EFL learners' prosodic distances in regard to PS:NS and SS:NS. The results are further visualized in Figure 5 below.

Table 6. Distances from the reference points

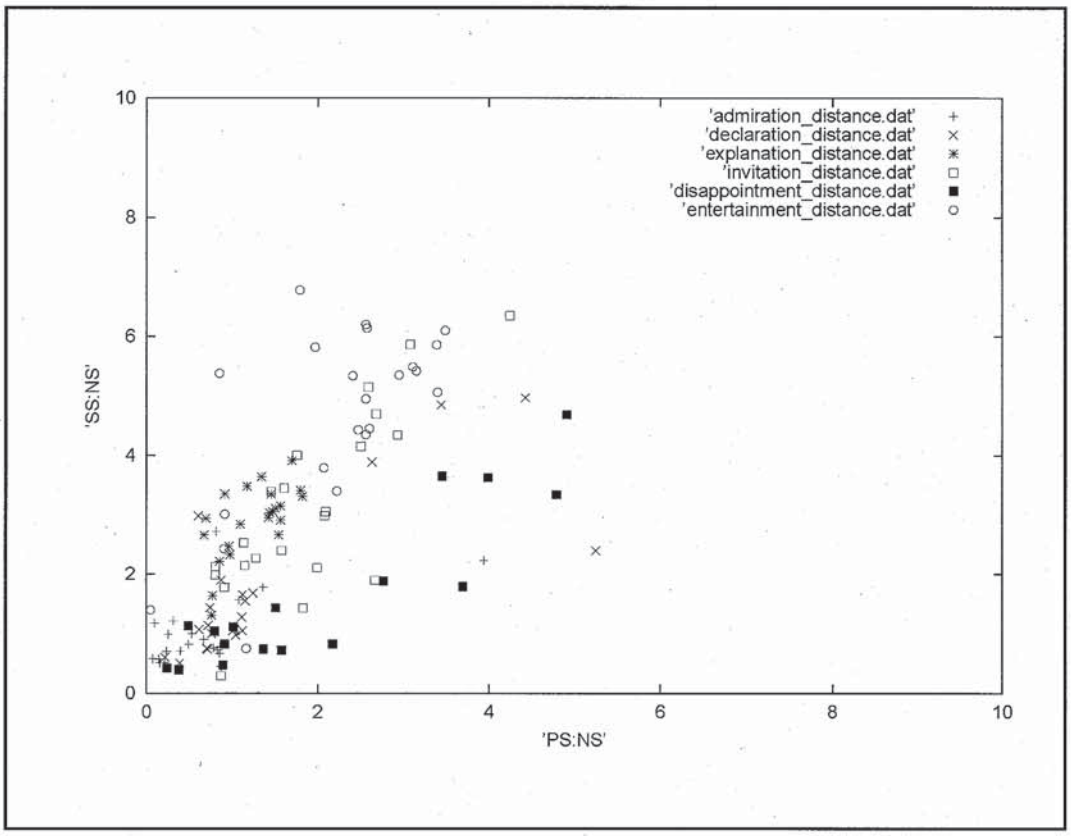
Items	PS:NS	SS:NS
	Mean	Mean
admiration	0.72	1.02
declaration	1.41	1.75
explanation	1.26	2.90
invitation	1.91	3.11
disappointment	1.92	1.80
entertainment	2.24	4.63
encouragement	1.37	
establishment	1.60	

Figure 5. Distances to the reference points



Notice that the PS:NS prosodic distances of all eight items are located between 0.5 and 2.5, while the SS:NS distances for the six items carrying secondary stress seem quite aberrant and exhibit little consistency. Among them, *admiration* shows the smallest distances on both PS:NS and SS:NS. *Disappointment* is unique in that it is the only item in which the PS:NS prosodic distance is greater than the SS:NS prosodic distance. The SS:NS prosodic distances of *explanation, invitation,* and *entertainment* are prominent, and the differences between their SS:NS and PS:NS distances stand evident. For all tokens of *admiration, declaration, explanation, invitation, disappointment,* and *entertainment,* the PS:NS prosodic distances and SS:NS prosodic distances are plotted on the horizontal axis and the vertical axis on the scatter chart below, respectively.

Figure 6. Prosodic distances to the points of references (learners)

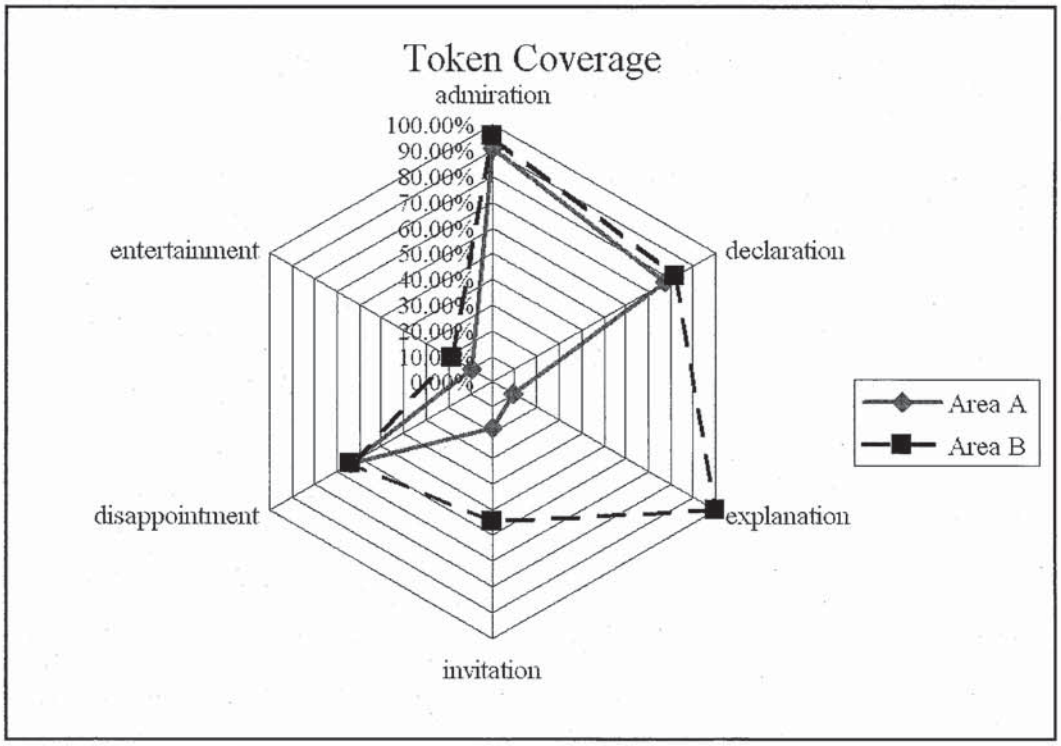


The chart suggests that distances are generally smaller on the PS:NS than on the SS:NS, and it seems that the tokens of each item share a particular area despite the fact that they are from different informants. In fact, we see that a significant number of dots are concentrated on the area of 0.0 to 2.0 on the PS:NS axis and 0.0 to 4.0 on the SS:NS axis. As a first approximation, two areas were selected as domains of sampling. The area A corresponds to the area of 0.0 to 2.0 on the PS:NS axis and 0.0 to 2.0 on the SS:NS axis; the area B corresponds to the area of 0.0 to 2.0 on the PS:NS axis and 0.0 to 4.0 on the SS:NS axis. Table 7 summarizes the coverage of the two respective areas by each of the six lexical items in question. The radar chart in Figure 7 visualizes the increased coverage from the area A to the area B.

Table 7. Token coverage

	Area A (x>2.0, y>2.0)		Area B (x>2.0, y>4.0)	
Items	# of tokens	frequency	# of tokens	frequency
admiration	20	90.91%	21	95.45%
declaration	17	77.27%	18	81.82%
explanation	2	9.09%	22	100.00%
invitation	4	18.18%	12	54.55%
disappointment	14	63.64%	14	63.64%
entertainment	2	9.09%	4	18.18%
Total	59	44.70%	91	68.94%

Figure 7. Token coverage in percentage



Of the 132 tokens (6x22=132), 59 fall within the area A, accounting for the coverage rate of 44.70%. The rate increases radically to nearly 70% in the area B with 91 tokens. What draws attention is the fact that coverage rates vary significantly item by item.

Consider the area A first. What strikes as most distinct is the fact that 20 out of the 22 tokens of *admiration* are located in this area, accounting for a coverage rate of 90.91%. This is immediately followed by *declaration* (77.27%) and *disappointment* (63.64%). The numbers of tokens of *invitation* (18.18%), *explanation* (9.09%), and *entertainment* (9.09%) remain extremely low.

Extending the sampling domain to the area B, we gain a different picture of coverage rates.

The area B contains 91 tokens and this accounts for 70%. It also holds 22 tokens of *explanation* despite its very small number of occurrences in the area A. Notably, these tokens are concentrated on the part of the area B that does not belong to the area A. *Invitation,* which has only four tokens in the area A, has increased up to 12, occupying a coverage of nearly 55%. It should be emphasized that there is not much increase in number between the areas A and B in regard to *entertainment* and *disappointment*. These results indicate that many of their tokens are located outside the area B. In fact, Figure 7 clearly shows the dots standing for the tokens of *entertainment* are concentrated in the upper region of the scatter chart and those for *disappointment* are scattered in a rather proportionate way between their PS:NS and SS:NS. Whether this is due to an intrinsic property of the suffix *-ment* or to some other factor is not clear at this stage.

5. Conclusion

The present study reported on some characteristics of word stress assignment that Japanese EFL learners reveal as a result of which class of suffix they choose in deriving a new word by suffixation. In so doing a method was proposed that measures one's performance on the realization of stressed syllables and their locations in the stated prosodic space. The results from the experiment conducted on a selected group of Japanese university EFL learners revealed that PS:NS ratios were distinctively high in compared with the other two dimensions whereas SS:NS ratios stand out with respect to the dimension of pitch.

Notes

This article is based on a paper read at the tenth Commemorative Conference of the 8th Phonetic society of Japan and the 1st International Congress of Phoneticians of English held at Kochi University on November 6, 2005.

1. In Ramus et al. (1999) a vocalic interval is referred to as the time span between the onset and the offset of a vowel or a vowel cluster and a consonant interval as the time span between the onset and the offset of a consonant or a consonant cluster.
2. The online edition of the *Exceed,* available at http://dictionary.goo.ne.jp/, contains approximately 120,000 entries.
3. WaveSurfer has been developed at the School of Computer Science and Communication of the Royal Institute of Technology in Stockholm. The software is downloadable from http://www.speech.kth.se/wavesurfer/

References

Abercrombie, D. (1967). *Elements of general phonetics.* Edinburgh: Edinburgh University Press.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

Galves, A., Garcia, J., Duarte, D., & Galves, C. (2002). Sonority as a basis for rhythmic class discrimination. Aix-en-Provence, France, 2002. Retrieved from *Speech Prosody 2002* at http://www.lpl.univ-aix.fr/sp2002/papers.htm, Université de Provence, Aix-en-Provence, France.

Jian, Nan. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*(1), 47-77.

Kiparsky, P. (1982). Lexical morphology and phonology. In The Linguistic Society of Korea (Ed.). *Linguistics in the morning calm.* (pp. 1-91). Seoul: Hanshin.

Laver, J. (1994). *Principles of English phonetics.* Cambridge: Cambridge University Press.

Mattys, S. L. (2000). The perception of primary and secondary stress in English. *Perception & Psychometrics, 62*(2), 253-265.

Mohanan, K. P. (1982). *Lexical phonology.* Unpublished doctoral dissertation, MIT.

Pike, K. L. (1945). *The intonation of American English.* Ann Arbor: University of Michigan Press.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265-292.

Reizevoort, J. & van Heuven, V. J. (1993). Pattern-driven morphological decomposition. Retrieved from the Digital Repository of University Leiden at https://academia.leidenuniv.nl/retrieve/3553/167_042.pdf.

Yamaoka, T. (2005). From item-learning to category-learning: A learning process of procedural knowledge of language. *Annual Review of English Language Education in Japan, 16*, 21-30.

（いおろい　たかひろ・本学助教授）