Asymmetry in Acquisition of Prosodic Structures: Evidence from L2 Japanese by English Speakers

Tokiko Okuma

1. Introduction

This study investigates the influence of L1 prosodic structures on the acquisition of simple and compound nouns in Japanese by L1 English speakers. Japanese simple and compound nouns have different prosody from their English counterparts. Japanese simple nouns can have no position of prominence (pitch fall), whereas English nouns always have prominence (stress). Japanese noun-noun compounds have one position of prominence as a whole, while English noun-noun compounds retain two positions of prominence. Under the assumption that Japanese pitch fall corresponds to feet (Shinohara, 2002), these prosodic differences are attributable to differences in prosodic organization between the two languages. If L2ers transfer L1 prosodic structures, L1 English learners of L2 Japanese should initially have a problem with Japanese simple and compound nouns because their L1 has different structures, as discussed in the following sections.

2. English stress and foot structure

2.1. English stress

English is a stress language with the main metrical parameters set as in (1). Each lexical word forms a prosodic word (PWd), which acts as the domain in which stress is computed. In a noun, the right-most syllable is extrametrical, and bimoraic trochaic feet are iteratively constructed from right to left. The position of stress is sensitive to rhyme structure. When the penultimate syllable is heavy, the stress falls on it; when the penultimate syllable is light, the stress falls on the antepenultimate syllable (Latin/English accent rule, Hayes, 1995, p.91). Stress is expressed through a combination of pitch, intensity and duration (Fry, 1955; Lieberman, 1960).

(1) English noun setting for metrical parameters

Foot Binarity: yes	Foot Iterativity: yes
Foot Shape: left	Extrametricality: yes
Foot Directionality: right-to-left	

2.2. Compound stress in English

In English compounds, each constituent retains stress. In a N_1N_2 compound, N_1 stress is elevated to main prominence status of the whole compound, and N_2 stress is downgraded to a secondary level (e.g. BLÁCK bòard)(stress subordination in Liberman & Prince, 1977). By contrast, in noun phrases, N_2 stress has main prominence and N_1 stress is subordinated (e.g. blàck BÓARD). Following the

* Tokiko Okuma, McGill University, tokiko.okuma@mail.mcgill.ca. I would like to thank Heather Goad and Lydia white for their guidance and support throughout this project. Thanks go to the GALANA 2012 audience and two anonymous reviewers. This research was supported by SSHRC grant #410-2011-0809 and FQRSC grant #2010-SE-130727 to Lydia White.

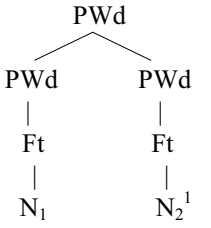
 Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012), ed. Chia-Ying Chu et al., 22-33. Somerville, MA: Cascadilla Proceedings Project.

prosodic hierarchy (e.g. Nespor & Vogel, 1986; Selkirk, 1986) in (2), N_1N_2 compound structures are given in (3). In (3), N_1 and N_2 form a PWd individually, and the two are conjoined into a higher PWd.

(2) Prosodic hierarchy (e.g. Nespor & Vogel, 1986; Selkirk, 1986)

Phonological Phrase (PPh)
|
Prosodic Word (PWd)
|
Foot (Ft)
|
Syllable (σ)

(3) English compounds



3. Japanese accent and foot structure (Tokyo Japanese)

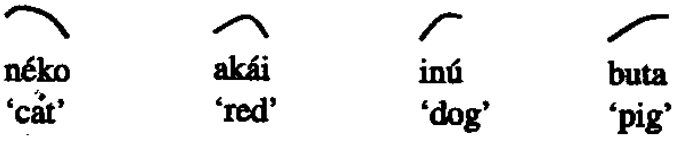
3.1. The tonal system

Japanese is a pitch accent language in which accents are realized by a drop in pitch. The pitch falls from high (H) to low (L) immediately after the accented mora. The accent (H tone) is lexically stored and can fall on any syllable. Tokyo Japanese is the dialect used as standard Japanese (Haraguchi, 2001) and classroom learners are exposed to. It has three rules affecting accent profile. First, once pitch falls, it remains L on all following morae in the word. That is, each word has at most one accent. Second, the pitch of the first mora of a word is L unless the mora has an accent (Initial Lowering Rule). Third, unaccented mono-moraic words are H by default. From the first two rules, 4 mora nouns can have 5 different pitch patterns (head-high, 2 mid-high, tail-high and flat patterns), as in Table 1. In this table, underlined H represents the location of accent. The first 4 patterns have an accent, while the last flat pattern does not. The difference between the tail high pattern and the flat pattern is typically manifested when followed by an accentually neutral particle, such as the nominative case marker *–ga*. The pitch of the tail high words falls immediately before the nominative case marker (e.g. ka.mi.na.rí-ga LHH<u>H</u>-L). By contrast, the pitch of the flat words remains H on the case marker (e.g. ka.mi.ga.ta-ga LHHH-H). Phrase-final position has little effect on lexical pitch—if a noun ends with H, it is still H in phrase-final position. (3) presents examples of pitch contours (Ota, 2003).

Table 1. Pitch pattern of 4 mora nouns (Haraguchi, 2001 with modifications)

	Pitch patterns		Example nouns	Accented or not
1	Head-high	<u>H</u> L LL	ká.ma.ki.ri ‘mantis’	Accented words
2	Mid-high1	L <u>H</u> L L	a.sá.ga.o ‘morning glory’	
3	Mid-high2	L H <u>H</u> L	a.o.zó.ra ‘blue sky’	
4	Tail high	L H H <u>H</u> (L)	ka.mi.na.rí (ga) ‘thunder’	
5	Flat	L H H H (H)	ka.mi.ga.ta (ga) ‘hair style’	Unaccented words

(3) Pitch contours of Japanese words in citation forms (Ota, 2003)



<u>H</u>L (head high) L<u>H</u>L (mid-high) L<u>H</u> (tail high) LH (flat)

3.2. Compound accent positions

As mentioned above, in English N_1N_2 compounds, stress is not altogether lost, indicating that the two PWds are preserved. By contrast, Japanese N_1N_2 compounds usually have, at most, one single

[1] This structure, in which Ft dominates N_1 or N_2, is used for convenience. When N_1 or N_2 is long, the entire noun will not be contained in Ft (e.g. *woman doctor* $[[(\text{ˈwʊmən})_{Ft}]_{PWd}\ [(\text{dɑk})_{Ft}\text{<tər>}]_{PWd}]_{PWd}$).

accent[2]. The position of the compound accent (CA) is largely determined by the length of N_2. The length and accent profile of N_1 does not matter (McCawley, 1968; Kubozono, 2008). Kubozono (2008) and Shinohara (2002) suggest that the CA patterns are categorized into two groups in (4) when N_2 is up to 4 morae. The example compounds in (4) are taken from Kubozono (2008) and Shinohara (2002), which are represented as K and S, respectively.

(4) Compound accent patterns (underlined parts represent accent positions)
I. N_2 is 2 morae or less
a. N_2 accent is preserved: LH…H-HL
pérusha + néko → perusha-néko‘Persian cat’ (K, S)
koosoku+básu → koosoku-básu‘highway bus’ (K)
b. Accent on the final syllable of N_1: LH…H-LL
ákita+inú → akitá-inu ‘Akita dog’ (K)
tinomí +ko → tinomí-go ‘infant’ (K)
c. Unaccented: LH...H-HH
sjákai+tóo → sjakai-too ‘Socialist Party’ (S)
akagi+jamá → akagi-jama ‘Akagi mountain’ (S)
II. N_2 is 3 or 4 morae
a. N_2 accent is preserved: LH…H-HLL, LH…H-HHLL
bíiti+báree → beeti-báree‘beach volleyball’ (K)
yámato+nadésiko → yamato-nadésiko‘Japanese lady’ (K, S)
b. Accent on the first syllable of N_2: LH…H-HLLL
áisu + koohíi → aisu-kóohii‘iced coffee’ (S)
minami + america → minami-ámerica‘South America’ (K)

3.3. Foot structure

I use the term ‘compound nouns’ to refer to nouns which consist of a modifier and a head, following Kubozono (1995). In Japanese, compound nouns have phonological, morphological, semantic, and syntactic characteristics which are distinct from ordinary noun phrases. In morphology, compound nouns appear without any coordinating morpheme. They are also orthographically expressed as one word. In semantics, they often have a unique meaning distinct from a combination of the meanings of N_1 and N_2. In syntax, it is usually impossible to insert an adjective, which only modifies N_2, between the two constituents (Kubozono, 1995). Note that every compound does not necessarily have all four characteristics. Here I will focus on N_1N_2 compounds and discuss their phonological characteristics, namely, CAs.

There is no consensus on Japanese foot formation among researchers (Kubozono, 1999, pp.57-58) for two reasons. First, there seems to be little evidence for/against foot formation for lexical words (Shinohara, 2000). Much of the evidence found so far is not for lexical words but for derived words, such as truncations of loanwords (Itô, 1990) and hypocoristic name shortenings (Poser, 1990). As a result, the literature discussing foot structure for lexical words is relatively limited. Second, Japanese feet do not contain an intensity contrast or a durational contrast, which trochaic feet or iambic feet tend to have, respectively (the Iambic-Trochaic Law, Hayes, 1995). In Japanese, prominence is predominantly realized by a pitch fall immediately after the accented mora, as we saw in 3.1. There is little use of intensity or duration to realize prominence (Beckman, 1986; Kondo, 2009). Each mora is pronounced with similar duration, irrespective of position of accents (Port, Dalby & O’Dell, 1987).

In this study, I adopt Shinohara’s (2000) assumption that “there is only one foot for the accentuation in a Japanese prosodic word” (p.77). She proposes that Japanese default accentuation is on the head of a bimoraic trochaic foot whose right edge is aligned with the left edge of the final syllable of the word, as in (5). Further, she suggests that when the N_2 accent of a compound coincides

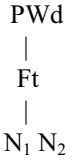
[2] When N_2 is further divided into multiple nouns, the compound accent is blocked and there can be multiple accents in the whole compound (Itô & Mester, 2003).

with the default accent, it is preserved as the CA, but when the N_2 accent is not default, the CA falls immediately preceding the morpheme boundary.

(5) Default accentuation: the head of the bimoraic trochaic foot (Shinohara, 2000)
(( ) indicates foot grouping, μ̲ represents accented mora.)
-(μ̲ μ) σ# (the final syllable σ can be light or heavy)

The default accent is observed in loanwords, proper nouns and in prosodically derived words, such as truncation and hypocoristics (Poser, 1990; Itô, 1990). Thus, Shinohara gives an account of CA formation, using Japanese default accentuation for non-compound words. Note that she admits exceptions. For example, she suggests that pattern in I-a of (4) is very limited in number, and treats it as an exception. Similarly, *tino(mí-go)*'infant', which ends with the foot with no following syllable, is another exception. In order to make *tino(mí-go)* compatible with (5), we need to add another assumption that 'the final syllable is not extra prosodic when N_2 consists of one mora' to (5). Thus, Shinohara's account may not be consistent with all compounds; nevertheless, it has two theoretical advantages over other proposals. First, it gives theoretical explanation for CA patterns, referring to default accents. Before Shinohara, there were descriptive generalizations for the CA patterns, including Kubozono (1995), but little theoretical explanation had been given. Second, she consistently uses one type of foot to explain compound accent location while other researchers, including Kubozono (1995, 2008), seem to use two (i.e. trochaic and iambic feet). Her analysis is not only simpler, but also more plausible than Kubozono's given that languages with two types of feet are either unattested or extremely rare. Based on this, I will adopt Shinohara's analysis in this study and assume that Japanese compounds have a structure in which N_1N_2 form one PWd, as in (6). (6) is entirely consistent with her analysis where the foot crosses over the N_1N_2 boundary.

(6) Japanese compound structure



4. Prosodic transfer and L2 acquisition

In L2 acquisition, it has been suggested that the initial state for L2 acquisition is the end state L1 grammar, and all abstract L1 properties can be transferred into the interlanguage grammar (the Full Transfer Full Access Hypothesis in Schwartz & Sprouse, 1996). In L2 acquisition of phonology, the acquisition task involves the changing of parameter settings (or constraint rankings, in constraint-based theories). For example, L1 English learners of L2 Japanese need to switch the parameter of Foot iteratively from Yes (English) to No (Japanese). The acquisition task may also need modification of L1 prosodic structures, such as projection of new structures and loss of structures. For example, L2ers from a L1 with CV syllables only must project a coda constituent when they learn English. Conversely, L2ers from a L1 which always needs feet must get rid of the foot as an obligatory constituent of the prosodic hierarchy when they learn Japanese. All of these are presumably not equal in difficulty.

Research on L2 acquisition of phonology has found that L2ers transfer aspects of the L1 for various levels of prosodic structure, such as syllable structure (Sato, 1984; Broselow & Finer, 1991) and foot structure (Archibald, 1998; Özçelik, 2011). However, most studies investigating word-internal prosody focus on morphologically simplex structures, such as simple nouns. As for morphologically complex structures, many studies, including Goad, White & Steele (2003), have examined the acquisition of functional categories. Acquisition of morphologically complex constructions of lexical categories has been understudied. This study investigates the influence of L1 prosodic structures on the acquisition of morphologically simplex and complex structures of lexical categories, namely, simple and compound nouns.

5. Research questions and predictions

As we saw in sections 2 and 3, Japanese and English prosodic structures differ in two respects. First, Japanese and English simple nouns are different in terms of the number of feet. Japanese simple nouns can be unaccented, and in this case, they have the structure without a foot in (7b). When accented, simple nouns have the structure with one foot in (7a). By contrast, in English, every PWd must contain a foot as in (7c), and feet are built iteratively. Second, Japanese and English compounds differ in terms of the number of PWds they contain. Japanese compounds form one PWd as a whole with one position of prominence as in (7d). This is shown from the accent pattern, crucially from the fact that the foot can cross over the compound boundary in Japanese. By contrast, English compounds form two PWds and have two positions of prominence as in (7e). In this study, it is hypothesized that L2ers initially transfer L1 prosodic structures. This leads to the predictions in (8).

(7)

Simple nouns			Compounds	
Japanese		English	Japanese	English
a. PWd | Ft | N (accented)	b. PWd | N (unaccented)	c. PWd | Ft | N (accented)	d. PWd | Ft | N_1 N_2 (either N_1 or N_2 is deaccented)	e. PWd PWd PWd | | Ft Ft | | N_1 N_2

(8) (i) L2ers will initially display the structure in (7c) for unaccented simple nouns, in place of (7b).
(ii) L2ers will initially display the structure in (7e) for compounds, in place of (7d).

(8i) suggests that L2ers will initially have a foot under every PWd, like (7c), failing to produce unaccented simple nouns in (7b). (8ii) suggests that L2ers will keep two PWds in N_1 and N_2, like (7e), failing to produce target-like compounds in (7d). They will produce two positions of prominence, which will be reflected as two pitch falls. Given that a previous study (Özçelik, 2011) suggests that getting rid of feet is difficult, besides, English has a handful of phonologically restructured compounds (e.g. postman is ['pous(t)mən]$_{PWd}$), (8i) would be more problematic than (ii) for L2ers.

6. The experiment

6.1. Informants

The informants were 10 L1 English learners of L2 Japanese (7 male and 3 female) and 6 native Japanese speakers (3 male and 3 female) as the control group. All L2ers started studying Japanese after 14 years of age, namely, after the supposed sensitive/critical period governing language development (Long, 1990). Before the experiment, the L2ers performed a story telling task to assess their oral proficiency. In the task, they described pictures in which a person was performing various activities. These pictures were used to encourage the informants to use a wide variety of accent patterns. Their readings were recorded and 2 native Japanese examiners rated their proficiency on a 5-point scale from 1 (= not at all native-like) to 5 (= perfectly native-like). The examiners had been working as Japanese language teachers in North America. The rating was based on two factors related to global accent: segmentals (i.e. how accurately each sound was pronounced) and fluency (i.e. how natural the intonation of phrases and sentences were). Based on the mean score of the ratings, 3 L2ers were advanced, 4 were intermediate, and the remaining 3 were elementary. The L2ers also performed a vocabulary check to confirm their knowledge of the 25 nouns used in the experiment. These nouns consisted of 1-4 morae and had different accent patterns (i.e. H, LH, HL, LHH, HLL, LHHH[3]). It was

[3] Other patterns, such as LHL and LHLL, were not included because they are limited in number (LHL and LHLL about 10% of all 3 and 4 mora nouns respectively) (Akinaga, 2011).

expected that the L2ers would know these nouns because they are common nouns taught in elementary Japanese language classes. One elementary subject, who got the lowest score 65% on the vocabulary check, was eliminated from the analysis because he could not fluently pronounce words that he said he knew. As a result, 9 L2ers, who knew at least 78% of the vocabulary items (90.3% on average) remained as subjects, as Table 2 shows. In the table, the L2ers are arranged in the order of the oral proficiency scores (i.e. E1 is the most advanced, while E9 is the least advanced).

Table 2. Individual profile of the L2ers

no	oral proficiency rating			vocab check (%)	age of onset	age	gender	formal instruction (years)	immersion in Japan (years)
	Proficiency group	mean score	individual score						
E1	advanced	3.1	3.3	100	20	31	M	4	2.3
E2			3.0	87	18	20	F	1.5	0
E3			3.0	96	16	20	M	2.5	0.8
E4	inter-mediate	2.6	2.8	100	24	33	M	0	6
E5			2.8	78	19	21	M	2.5	0
E6			2.5	87	14	20	F	5	0.8
E7			2.5	87	17	20	F	1.5	0.3
E8	elementary	1.3	1.5	78	14	19	M	5.5	0
E9			1.0	100	33	43	M	3	5

6.2. Materials and stimuli

The informants were shown three simple nouns with pictures (see (9a)) and pronounced them in isolation. This was to see if they were able to pronounce the nouns with correct accents. In the stimuli, all words were written in Japanese without English translations, and accents are not orthographically written in Japanese. Next, they produced a compound by combining the two nouns given, looking at pictures, as in (9b). Finally, they pronounced the compound in a Japanese carrier sentence 'Here is a ….' as in (9c). This was to avoid transfer of phrase final lengthening and pitch fall in English to Japanese. The compounds were either novel (n=32) or real (n=8). The novel compounds were typically names of things in the shape of animals, consisting of 4-6 morae with different accent patterns (i.e. LHH-LL, LHH-HLL, L-HLL). They were divided into 8 conditions, depending on the N_1 and N_2 accents, as Table 3 shows. The real compounds represented the same 8 conditions (e.g. Condition 1: *kabutó-musi* (LHH-LL) 'beetle', consisting of *kábuto* (HLL) 'armor' and *mu.si* (LH) 'insect', Condition 2: *Kyootó-eki* (LHH-LL) 'Kyoto Station', consisting of *Kyóoto* (HLL) 'Kyoto' and *éki* (HL) 'station'). In this study, accentuation of real words follows Akinaga (2011), one of the two major accent dictionaries of Tokyo Japanese. The novel and real compounds were randomly presented to encourage the informants to pronounce the novel compound as one word, just like real compounds. The stimuli were presented to the informants on a computer screen in a sound-attenuated lab. Their oral production was recorded in Praat (Boersma & Weenink, 2011) using a head mounted microphone. After the recording, the simple nouns and compounds were transcribed, and the position of accents was determined through examination of the fall of the fundamental frequency (F0) contour in Praat by the author and another native Japanese judge. The data the two judges disagreed on in deciding the accent positions were eliminated from the analysis. Mora boundaries were annotated in text grid files, and the F0 measures were extracted. Intensity and duration measures were extracted for some subjects.

(9) Stimuli example

a.

ryokoo 'traveling'

pan 'bread'

usage 'rabbit'

b.

usagi 'rabbit' + pan 'bread' → (?)

c. The carrier sentence

(___?___)-ga ari-masu. 'Here is a (?)'
-Nom is

Table 3. Stimuli (novel compounds, n=4×8=32)

cond	CA patterns	examples	N_1 (1 or 3 morae)		N_2 (2 or 3 morae)	
1	LHH-LL	pandá-futa 'panda-shaped lid'	HLL	pá.n.da 'panda'	LH	fu.ta 'lid'
				ká.ra.su 'crow'		mo.ti 'rice cake'
2		pandá-pan 'panda-shaped bread'	HLL	pá.n.da 'panda'	HL	pá.n 'bread'
				ká.ra.su 'crow'		má.do 'window'
3		usagí-futa 'rabbit-shaped lid'	LHH	u.sa.gi 'rabbit'	LH	fu.ta 'lid'
				ne.zu.mi 'mouse'		mo.ti 'rice cake'
4		usagí-pan 'rabbit-shaped bread'	LHH	u.sa.gi 'rabbit'	HL	pá.n 'bread'
				ne.zu.mi 'mouse'		má.do 'window'
5	LHH-HLL	panda-zúbon 'panda-print pants'	HLL	pá.n.da 'panda'	HLL	zú.bo.n/zu.bó.n 'pants'
				ká.ra.su 'crow'		mé.ga.ne 'eye glasses'
6		panda-tókee 'panda-shaped clock'	HLL	pá.n.da 'panda'	LHH	to.ke.e 'clock'
				ká.ra.su 'crow'		ka.ba.n 'bag'
7	L-HLL	ka-zúbon 'mosquito-print pants'	H	ka 'mosquito'	HLL	zú.bo.n 'pants'
				ha 'leaf'		mé.ga.ne 'eye glasses'
8		ka-tókee 'mosquito-shaped clock'	H	ka 'mosquito'	LHH	to.ke.e 'clock'
				ha 'leaf'		ka.ba.n 'bag'

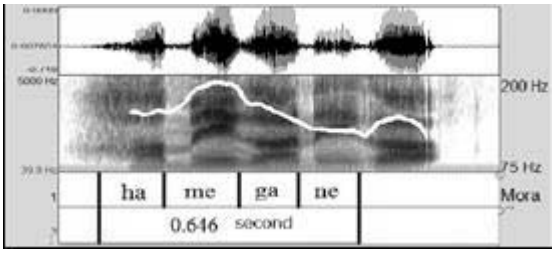
7. Results

7.1. Acoustic realization of accents

This section presents how the informants produced accents through pitch (F0) measurements because Japanese accents are realized by a sudden pitch fall after the accented mora, not duration or intensity, as we saw in 3. (10) shows examples of a novel compound (*ka-mé.ga.ne* LHLL 'mosquito-shaped eye glasses') by the male native Japanese speaker (J1) and the male L2er (E9). The white curves on the spectrograms represent the pitch. (10a) shows that the pitch goes up in the second mora and hits peak (191Hz), then drops to 134Hz in the third mora. This pitch fall is perceived as an accent on the second mora. By contrast, there is no drastic pitch fall in (10b). All morae stay in between 95 and 114 Hz except at the very end of the 4th mora. Therefore, no accent is perceived. Tables 4 and 5 present pitch range (i.e. the difference between the maximum F0 and the minimum F0 in the word) by the native Japanese speakers and the L2ers. Two findings are shown in these tables. First, the average of pitch range of the words by the L2ers (43-53 Hz) is 72-79% of the controls (59-67 Hz). This is probably due to L1 transfer because the L2ers also produced English compounds (*guitar string*, *White House*, *English teacher*, *blackboard*), and the mean pitch range of the English compounds was 30.8 Hz. Second, the pitch drop ratio following the accented mora by the native Japanese speakers was 44-48%. By contrast, the pitch drop in the same mora by the L2ers was 27-30%, which was significantly smaller than that of the native Japanese speakers (Mann-Whitney U Test, U=1.0, p=.002<.01).

(10) Examples of wave forms

a. The native Japanese speaker (J1)



b. The L2er (E9)

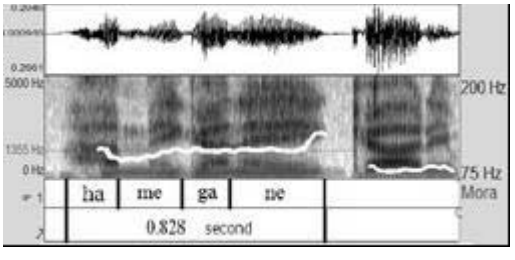


Table 4. Pitch results by the native Japanese speakers (Hz)

	mean			individual					
	range	drop after accent[4]	drop ratio (%)	range (male)			range (female)		
				J1	J2	J3	J4	J5	J6
Novel words	59	28	0.48	65	38	42	70	77	61
Real words	67	29	0.44	59	29	51	71	95	94

Table 5. Pitch results by the L2ers (Hz)

	mean			individual								
	range	drop after accent	drop ratio (%)	range (male)						range (female)		
				E1	E3	E4	E5	E8	E9	E2	E6	E7
Novel words	43	15	0.30	53	19	64	35	30	26	42	65	55
Real words	53	16	0.27	58	24	81	36	31	44	53	84	69

7.2. Accuracy of accent position (simple nouns)

Table 6 presents the accuracy in producing accents on simple nouns. 10 accented nouns and 13 unaccented nouns were targeted in total; of these, only the nouns which the L2ers knew in the vocabulary check were analyzed. The number of tokens of each noun varied from 1 to 9, and there were maximally 78 tokens for one informant (39 tokens each for the accented nouns and the unaccented nouns). Table 6 shows that the L2ers were divided into three proficiency groups; nevertheless, there was little difference between the groups; the group accuracy was 52% for the advanced (E1-E3), 60% for the intermediate (E4-E7), 47% for the elementary, and 54% on average for all groups (range 35-74%). There was no strong correlation between accuracy and proficiency (r=.28). In the table, the 'Accented' row presents the accuracy of accented simple nouns, and the 'Un-accented' row presents the accuracy of the un-accented simple nouns. The L2ers were generally more accurate in producing the unaccented nouns (81% on average) than the accented nouns (29% on average). Especially, 3 out of the 9 L2ers (E3, E5, and F7) were 100% accurate with the un-accented nouns.

Table 6. Accuracy for simple nouns (%)

	E1	E2	E3	E4	E5	E6	E7	E8	E9	mean
Accented	59	0	3	62	18	95	18	0	0	29
Un-accented	85	77	100	44	100	47	100	97	82	81
mean	72	35	49	53	52	74	59	53	41	54
group mean	52 (advanced)			60 (intermediate)				47 (elementary)		

7.3. Accuracy of accent position (novel compounds)

Table 7 shows accuracy of accent positions for the novel compounds. Note that the L2ers produced them as one words (i.e. compounds), not just a string of two words. The tokens were the second production of the compound (see (9b) and (9c)); in addition, they were produced in a carrier sentence. Moreover, tokens with a pause on N_1N_2 boundary were treated as an error (1 token by E5).

Table 7 shows that the mean accuracy was 22%, and accuracy by the individual L2er varies from 4 to 54%. Note that the low accuracy does not seem to be attributable to the misunderstanding of the N_1 or N_2 accents because the CA positions are determined by the number of mora in N_2, not the accent patterns of N_1 or N_2. Even if they misunderstood the accentuation of N_1 or N_2, they should have produced the same compound accents as native Japanese speakers if they knew the rule, which places the CA on the trochaic foot. Therefore, the low accuracy results suggest that the L2ers potentially created unique rules. Table 7 also shows the distribution of the non-target like accents on the novel

[4] 'drop after accent' column in Table 4 shows the pitch fall on the mora which immediately follows accented mora. 'drop after accent' column in Table 5 shows the pitch fall on the same mora as that column in Table 4.

compounds. There are two findings. First, the L2ers did not produce multiple accents in a compound. That is, all compounds were either unaccented or had one accent. This suggests that the L2ers knew that Japanese compounds have only one foot, unlike English compounds. Second, there was variation in the error patterns. The L2ers are divided into 4 groups by their error patterns: (i) the N_1 accent group, (ii) the N_2 accent group, (iii) the unaccented group[5], and (iv) no pattern group.

Table 7. Accuracy and error patterns on the novel compounds (%)

			(i) N_1 accent group				(ii)N_2 accent group		(iii) Un-accented group		(iv)No pattern	mean
			E1	E4	E6	E7	E2	E8	E3	E9	E5	
accurate			13	25	54	17	8	28	4	28	29	22
not accurate			87	75	46	83	92	72	96	72	71	78
Error patterns	Wrong accent on	N_1	65	75	42	46	17	16	0	16	23	36
		N_2	18	0	4	8	59	45	0	6	20	16
	No accent		4	0	0	29	16	11	96	50	23	26
	separated nouns		0	0	0	0	0	0	0	0	5	0

8. Discussion

8.1. L2 prosodic structures

In section 5, I made following predictions,

(i) L2ers will initially display the structure in (7c) for unaccented simple nouns, in place of (7b).

(ii) L2ers will initially display the structure in (7e) for compounds, in place of (7d).

Prediction (i) was not supported. The accuracy for unaccented simple nouns was 81% on average, suggesting that the structure in (7b) was not problematic. By contrast, the accuracy for accented simple nouns was 29%. The error analysis shows that the L2ers overused unaccented nouns in place of accented nouns. As far as we rely on the pitch analysis, this result suggests that the L2ers produced target-like simple nouns without feet in (7b). Prediction (ii) was not confirmed, either. No L2er produced compounds with two accents, as presented in 7.3. This suggests that the L2ers did not use the L1 structure in (7e). Note that the low accuracy of the CA positions (e.g. 22% for the novel compounds) does not contradict the successful acquisition of the structure in (7d). The L2ers wrongly put the CA on N_1 or N_2 52% of the time, nevertheless, they produced one accent, not two accents, in the compound. This means that they produced target-like compounds with one PWd in (7d) 74% of the time though the position of the accent was not necessarily accurate. They also produced unaccented compounds (26% of the time). This suggests that they produced footless compounds. To summarize, the pitch analysis shows that the L2ers produced the four structures in (11).

(11) L2 structures

Simple nouns		Compounds	
a. target-like	b. target-like	c. target-like	d. non-target-like
PWd	PWd	PWd	PWd
|	|	|	|
Ft	N	Ft	N_1N_2
|	(unaccented)	|	(unaccented)
N		N_1 N_2	
(accented)		(accented)	
Accuracy 29%	81%	production rate 74%	26%

[5] The accuracy by E7, E8, and E9 were similar, but their error patterns were significantly different (e.g. E7 vs. E8: $\chi^2(2, N=155)=47.91, p<.01$).

8.2. Intensity analysis

The pitch analysis so far suggests that the L2ers produced PWds without feet. However, it may be too early to jump to the conclusion that the L2ers have successfully eliminated feet. In order to clarify whether the L2ers indeed produced target-like structures, we need to look at intensity, the acoustic measurement corresponding to the loudness of a sound (Ladefoged, 2001), for two reasons. First, English stress correlates not only include pitch but also intensity. The L2ers may have divorced pitch from foot structure, which would make them sound more target-like, but continued to use intensity to mark foot structure, thereby respecting L1 prosodic constraints. If this is the case, we may still find evidence of feet when we look at intensity. Second, the L2ers had a small pitch range, as we saw in Tables 4 and 5. This may suggest that pitch may not be a strong cue to realize prominence for them. Recall that intensity is not an acoustic correlate of accent in Japanese. When a native control produced compounds, the first or second syllable tended to have higher intensity, irrespective of accent positions. For these reasons, the remaining part of this section investigates (i) whether the L2ers truly got rid of feet when they produced unaccented simple nouns and unaccented novel compounds and (ii) whether the L2ers produced one PWd or two PWds when they produced unaccented novel compounds, by looking at intensity values for E3, E5, E7, and E9. These subjects are under focus because they produced unaccented compounds more than 23% of the time[6].

Table 8 and Table 9 show examples of the unaccented productions by E9, in which feet are observed by the intensity measurements. The tables contain only 4 productions for limitation of space. The tables show the maximum intensity of each syllable of the simple nouns and novel compounds which were produced without accents. The head of the foot (i.e. the most intense syllable in the word) is bold, and the right-most column presents the difference in intensity between the head and the adjacent syllable. When the head is more than 5.5 dB higher than the adjacent syllable, they are considered to form a foot. The criterion value, 5.5 dB, is adopted for two reasons. First, when E9 pronounced English nouns (*potáto* and *dóctor*), the intensity of the accented syllable was 5.5 dB higher than the following syllable on average. Second, Fry (1955) found that people judge two syllable verbs (e.g. *permít*) as a noun (i.e. *pérmit*) more than 60% of the time when the first syllable of two syllable verbs is pronounced more than 5 dB stronger than the second syllable. By this criterion, 15 out of total 35 unaccented productions of simple nouns (i.e. 43%) and 7 productions out of total 16 unaccented compounds (i.e. 44%) produced by E9 are considered as evidence for forming left-headed binary feet. Intensity of the head (mean 70.8, range 67.9-73.8 dB) is significantly larger than the non-head syllable (mean 63.3, range 59.5-67.1 dB) in the foot (Mann-Whitney U Test, U=0, p<.001). The difference of intensity between the head syllable and the non-head syllable in the foot (mean 7.5, range 5.6-10.4 dB) is significantly larger than the difference in intensity between the most intense syllable and the following syllable (mean 2.4, range 0.0-5.3 dB) in the remaining productions, in which feet are not observed (Mann-Whitney U Test, U=0, p=<.001). Thus, E9 built moraic trochaic feet at least 43% of the time. This suggests that E9 did not get rid of feet, contrary to the pitch analysis. He separated pitch from intensity, and the former was target-like, while the latter was not.

Table 8. Intensity (simple nouns) (dB)

words	1 syll	2 syll	
zu(bón)	61.2	**68.7**	7.5
(má.do)	**73.8**	63.5	10.4
to(kée)	60.2	**70.6**	10.5
fu(tá:)	58.7	**71.5**	12.9

Table 9. Intensity (novel compounds) (dB)

words	1 syll	2 syll	3 syll	4syll	
pan(dá-mo)ti	66.4	**70.5**	62.4	65.1	8.0
(pán)(dá-to)kee	**70.3**	**71.0**	63.1	68.0	7.9
(ká-zu)bon	**72.7**	67.1	67.1		5.6
(ká-to)kee	**69.3**	59.5	63.0		9.8

Table 9 suggests that E9 produced trochaic feet in novel compounds. What is crucial here is that the feet cross over the N_1N_2 boundary (e.g. *(ká-zu)bon*). This suggests that he successfully produced one PWd, not two PWds, consistent with the pitch analysis. Duration measurements also suggest that

[6] The remaining L2ers (E2 and E8) also produced unaccented compounds but the frequency was relatively limited (16% and 11% of the time).

he produced one PWd. The vowel duration of mono-moraic N_1 (i.e. the initial vowel in *(ká-to)kee* and *(ká-zu)bon* is 0.079ms on average, which is not significantly longer than the same vowel of longer N_1 (e.g. the initial vowel in *karasu-pan*, 0.076ms on average). If he had produced two PWds, the vowel of the mono-moraic N_1 would have lengthened to satisfy Foot Binarity (e.g. *(ká:)-to(kée)*), however, this is not observed. Thus, intensity and duration measurements show that E9 produced target-like compounds with one PWd. As for foot iterativity, there is no conclusive evidence. To sum up, in contrast to the pitch analysis, which suggested the structures in (11), the intensity analysis indicates that E9 has the non-target-like structures with feet.

Intensity measurements reveal that E5 produced similar feet to E9, suggesting that E5 has the non-target-like structures. In contrast to E9 and E5, intensity measurements do not indicate the presence of feet in speech of E7 and E3. As for E7, the first syllable consistently had the highest intensity, irrespective of the number of syllables in a word and noun type (i.e. simple noun vs. compound). The intensity differences between the 1st and 2nd syllables are 2.9 dB on average (range 0.6-5.4 dB). As for E3, positions of syllables with the highest intensity are randomly located throughout the word.

To summarize, the intensity analysis shows that E9 and E5 produced trochaic feet when they produced simple nouns and compounds with no pitch fall. This suggests that they did not get rid of feet in simple nouns and compounds, contrary to what the pitch analysis suggested. E9 and E5 separated pitch from intensity, and they successfully used the former to pronounce unaccented words, sounding more target-like, nevertheless, the latter shows the existence of feet. By contrast, their feet in compounds crossed over the N_1N_2 boundary. This suggests that they successfully produced one PWd as a whole, not two PWds, in a compound. This result is consistent with the pitch analysis, which suggests that the L2ers formed one PWd in accented compounds. From these results, it can be concluded that modification of existing prosodic constituents, PWds, is acquirable at early stages of development, whereas elimination of existing prosodic constituents (feet) is more problematic. This conclusion provides some support for Özçelik (2011), who suggests that once a prosodic constituent is projected in the L1, it is impossible to eliminate it from the L2 grammar. However, most of the L2ers in this study were residents of Montreal, studying Japanese as a foreign language rather than a second language. If they had more natural input, it could be possible for them to get rid of feet, like E3.

8.3. Limitations of the study

There are three shortcomings in this study. First, the number of informants was limited. More informants need to be tested to make the results more reliable. Second, this study only tested L1 English speakers. L2ers whose L1s are not English should be tested to see whether or not elimination of prosodic constituents is truly more problematic than modification of prosodic structures. Finally, this study only investigated the production ability by L2ers. Whether or not L2ers perceive the pitch as native Japanese speakers do, should be investigated in future research.

9. Conclusion

This study investigated whether L1 English learners of L2 Japanese successfully acquire the prosodic structures of Japanese simple and compound nouns. Under the assumption that Japanese pitch accents correspond to the trochaic feet (Shinohara, 2000), Japanese unaccented simple nouns are footless. Japanese compounds form one PWd as a whole with one foot that may span the compound boundary. These structures differ from English simple and compound nouns; as a result, the L2ers were expected to have problems with these structures if they transfer L1 prosodic structures. In the study, 9 L2ers were compared with native Japanese speakers in producing real simple nouns and novel compound nouns, and their pitch and intensity were measured. The pitch analysis suggested that the L2ers produced target-like simple nouns without feet at an accuracy rate of 81%. However, the intensity analysis revealed that at least two L2ers were not target-like, producing unaccented nouns with L1 feet. As for compounds, both the pitch and intensity analysis suggested that they successfully produced one PWd, not two PWds. These results suggest that modification of existing prosodic structures, namely, PWds, is acquirable at early stages of development, whereas elimination of existing prosodic constituents, namely, feet, is more problematic.

References

Akinaga, Kazue (2011). *Shinmeikai Nihongo Akusento Jiten* [Shinmeikai Japanese accent dictionary]. Tokyo: Sanseido.

Archibald, John (1998). *Second language phonology.* Philadelphia: John Benjamins.

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

Boersma, Paul & Weenink, David (2011). *Praat: Doing phonetics by computer*. [Computer Software] Amsterdam: Department of Language and Literature, University of Amsterdam. Retrieved from http://www.praat.org/

Broselow, Ellen & Finer, Daniel (1991). Parameter setting in second language phonology and syntax. *Second Language Research 7*, 35-59.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America 27*, 765-768.

Goad, Heather, White, Lydia & Steele, Jeffery. (2003). Missing inflection in L2 acquisition: Defective syntax or L1-constrainted prosodic representations? *Canadian Journal of Linguistics 48 (3/4)*, 243-263.

Haraguchi, Shosuke (2001). Accent. In Natsuko Tsujimura (Ed.) *The handbook of Japanese linguistics*. Oxford: Blackwell. Blackwell Reference Online. 16 June 2011.

Hayes, Bruce (1995). *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press.

Itô, Junko (1990). Prosodic minimality in Japanese. In *CLS 26-II; Papers from the parasesssion on the syllable in phonetics and phonology*, 213-239.

Itô, Junko & Mester, Armin (2003). *Japanese Morphophonemics: Markedness and Word Structure*. Cambridge, MA: MIT Press.

Kondo, Mariko (2009). Is acquisition of L2 phonemes difficult? Production of English stress by Japanese speakers. In Melissa Bowles, Tania Ionin, Silvina Montrul & Annie Tremblay (Eds.) *Proceedings of the 10th Generative Approaches to Second Language Acquisition Conference (GASLA, 2009),* 105-112. Somerville, MA: Cascadilla Press.

Kubozono, Haruo (1995). *Gokeisei-to oninkouzou* [Word formation and phonological structures]. Tokyo: Kuroshio Publishers.

Kubozono, Haruo (1999). Mora and syllable. In N. Tsujimura (Ed.) *The handbook of Japanese linguistics*. Oxford: Blackwell Publishing.

Kubozono, Haruo (2008). Japanese accent. In Shigeru Miyagawa & Mamoru Saito (Eds.) *The Oxford handbook of Japanese linguistics*, 165-191. Oxford: Oxford University Press.

Ladefoged, Peter (2001). *A course in phonetics* (4th edition). Boston: Heinle & Heinle.

Liberman, Mark & Prince, Alan (1977). On stress and linguistic rhythm. *Linguistic Inquiry 8*, 249-336.

Lieberman, Philip (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America 32,* 451-454.

Long, Michel (1990). Maturational constraints on language development. *Studies in Second Language Acquisition 12*, 251-285.

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

Nespor, Marina & Vogel, Irene (1986). *Prosodic phonology*. Dordrecht: Foris.

Ota, Mits (2003). The development of lexical pitch accent systems: An autosegmental analysis. *Canadian Journal of Linguistics 48 (3/4)*, 357-383.

Özçelik, Öner (2011). *Representation and acquisition of stress: The prosodic subset principle and the case of Turkish*. PhD dissertation. McGill University.

Port, Robert, Dalby, Jonathan& O'Dell, Michael (1987). Evidence for mora timing in Japanese. *Journal of the Acoustic Society of America 81*, 1574-1585.

Poser, William (1990). Evidence for foot structure in Japanese. *Language 66*, *1*, 78-105.

Sato, Charlie (1984). Phonological processes in second language acquisition: Another look at interlanguage syllable structure. *Language Learning 37*, 247-272.

Schwartz, Bonnie & Sprouse, Rex (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research 12*, 40-72.

Selkirk, Elisabeth (1986). On derived domains in sentence phonology. *Phonology3*, 371-405.

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

Shinohara, Shigeko (2002). Metrical constraints and word identity in Japanese compound nouns. *MIT working papers in Linguistics 42*, 311-328.

Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)

edited by Chia-Ying Chu,
Caitlin E. Coughlin, Beatriz Lopez Prego,
Utako Minai, and Annie Tremblay

Cascadilla Proceedings Project Somerville, MA 2014

ISBN 978-1-57473-463-8 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Okuma, Tokiko. 2014. Asymmetry in Acquisition of Prosodic Structures: Evidence from L2 Japanese by English Speakers. In *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, ed. Chia-Ying Chu et al., 22-33. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3063.