

Analysis of Japanese Pitch Accent in L1 and L2 Speech*

Naomi Ogasawara
University of Aizu

This study examines pitch patterns in Japanese speech produced by native speakers and Taiwan Mandarin speakers learning Japanese at the beginning level. It aims to provide detailed characteristics of pitch patterns at both word and sentence levels, and to discern common pitch errors in Mandarin speakers' pronunciations. Mandarin speakers had more difficulty with pronunciation of words consisting of all light syllables than with words containing a heavy syllable, and a steep pitch drop all the way through the words was commonly observed. More importantly, the learners seemed unaware of the difference between accented rising pitch followed by a low pitch and accent-less rising pitch followed by a high pitch. This suggests that even though they know how to correctly pronounce words, they need to know accurate pronunciations at phrase level.

Key words: pitch accent, L2 acquisition, Japanese, Taiwan Mandarin

1. Introduction

This study investigates pitch patterns in Japanese speech produced by native speakers of the Tokyo dialect and by non-native speakers learning Japanese as a foreign language at the beginning level. The target group of learners was Taiwanese students whose native language was Taiwan Mandarin. It is understandable to expect that Mandarin speakers would be keenly aware of pitch change in an L2 since their native language is a tone language. It is also reasonable to anticipate that the transfer of Mandarin tone contours would have some effect on pitch patterns when they speak Japanese. The study first tries to provide a phonetic analysis of pitch contours in L1 speech, including more detailed information than the general descriptions of pitch accent found in regular language textbooks. The study then tries to detect differences in pitch patterns in the learners' speech in comparison with those in native speech and tries to discern where the difficulty for the learners lies.

According to a survey conducted by the Japan Foundation, the number of learners of Japanese all over the world has been increasing in the last three decades (Japan Foundation 2013). The biggest group of learners is Mandarin speakers. In Taiwan, where Japanese is the second most popular foreign language after English, and there were 233,417 learners in 2012 (Japan Foundation 2014). In China, the number of learners reached over a million in 2012. Most learners' interests and motivations for learning Japanese is 'communication in Japanese' as well as

* I would like to thank Dr. Timothy Vance at the National Institute for Japanese Language and Linguistics for his insightful comments. I also thank the participants in these experiments at National Taiwan Normal University.

‘Japanese pop-culture, such as anime, manga, and fashion’ and ‘job hunting’ (Japan Foundation 2013, Hirano 2014).

Hirano (2014) pointed out that learning not only grammar but also good pronunciation is essential to meeting learners’ needs in improving their communication skills; however, there has not been much focus on teaching pronunciation in classrooms. Needless to say, accurate pronunciation of phonemes is one of the most important things, together with vocabulary and grammar that, learners should master, but in-class practice is not sufficient for acquiring ‘good’ pronunciation involving not only phonemic but also prosodic features.

Many cross-linguistic studies have discovered that native and non-native listeners consider the degree of prosodic distortion for judging foreign-accentedness of L2 speech (Anderson-Hsieh, Johnson & Koehler 1992, Jilka 2000, Ishihara et al. 2011, among others). Jilka (2000) examined perception of L2 German and L2 American English speech and found that distortion of intonation contributed most to hearing foreign accents besides contortion of segmental or other prosodic features. Ishihara et al. (2011) investigated adverse effects of timing and pitch errors on perception. Japanese sentences produced by native English speakers, who had received 160 hours of Japanese training, were used as stimuli in their study. The authors determined that compared with pitch errors, timing errors in long/short vowels or consonants (e.g. *kookoo* ‘high school’ and *chotto* ‘a bit’) were evaluated more harshly by native and non-native listeners whose native language was English, Chinese, or Bahasa Indonesian. The authors claimed that timing features are salient in Japanese, and learners could be learning these features earlier in the course of learning. Because the notation in *kana* writing is different for long and short segments, once learners are introduced to the *kana* writing system, they can become more aware of long/short contrasts in their pronunciation and acquire them more easily than pitch accents.

The findings from the study of Ishihara et al. (2011) suggest that learners and listeners consider pitch errors to be less problematic in non-native Japanese speech. Although pitch accents of words are lexical in Japanese, there are not so many word pairs contrasting in pitch accent patterns (e.g. /a¹me/ ‘rain’ and /ame/ ‘candy’); henceforth pitch errors may cause misunderstanding less frequently than phonemic errors. Munro & Derwing (1995) argued that foreign accent does not necessarily lower comprehensibility and intelligibility of L2 speech. Consequently, learners may pay less attention to pitch when they speak, such that correct pitch contours will be acquired eventually but perhaps not at an early stage of learning.

¹ The symbol “¹” indicates the syllable before the symbol has an accent.

Actually, most Japanese textbooks for beginners only provide a brief introduction to the pitch accent system in a couple of pages in a book, and insufficient hours in Japanese classes are devoted to teaching pronunciation (Shport 2008). This must be another reason for learners' lack of awareness and late acquisition of pitch accent. Hasegawa (1995) also noted some discrepancy between general descriptions about pitch accent in textbooks and actual pronunciations by native speakers. She pointed out that devoiced high vowels /i, u/² (e.g. *kita* 'north' and *desu* 'Copula-Be') cannot bear a pitch, such that a pitch-less syllable cannot contribute to forming a pitch contour of a word; however, there is no explanation for why pitch is assigned to a devoiced vowel. Another inconsistency is that the fundamental frequency (F0) peak in native speech often occurs after the syllable where a lexical accent is supposed to fall – this phenomenon is known as *oso-sagari* 'delayed F0 fall' (Neustupný 1966). However, Sugito (1970) and Hata & Hasegawa (1988) argued that native speakers still perceive an accent on the expected syllable when pitch drop follows the syllable. This phenomenon indicates some gaps between conventional descriptions of pitch accent and what is happening in actual speech. To fill such gaps for learners, Japanese language instructors should obtain good knowledge about pitch accent and incorporate this into their classroom teaching. As background, the following section summarizes the pitch accent system in Tokyo Japanese mostly discussed in the linguistic literature, not in Japanese textbooks.

1.1 Overview of the pitch accent system in Tokyo Japanese

As is well known, Japanese is a pitch accent language (Kubozono 2008, 2012, Vance 2008). Each mora³ example, /a^hme/ (HL) 'rain', /tama^hgo/ (LHL) 'egg', and /hac^higatsu/ (LHHH) 'August'. High pitch and low pitch are not defined as the absolute pitch height of pitch; rather, they are the relative height of pitch compared with the pitch level in adjacent moras (Kindaichi 1966). An accent is marked on the last high-pitched mora followed by a relatively low-pitched mora. Although there are many dialects of Japanese with different pitch accent systems (see Hirayama 1966 for details), describing dialectal accent systems is not our focus here. We focus on the pitch accent system of the Tokyo dialect in this paper since it is the most commonly learned dialect by non-native speakers.

The number of pitch contours in nouns is systematic, following the 'n+1 rule'

² High vowels in Japanese are usually devoiced between voiceless consonants or between a voiceless consonant and a pause.

³ The number of moras and syllables are the same for light syllables. A heavy syllable carries two moras; for example, [tombo] has two syllables but three moras. It is thought that moras carry pitch but syllables bear accent.

(Akinaga 1966, Kubozono 2008, 2012) where ‘n’ represents the number of syllables in a word. For example, there are three pitch patterns found in dimoraic words: HL (e.g. /a¹me/ ‘rain’), accented LH (e.g. /jama¹ ga (L)/ ‘mountain + nominative case particle (NOM)’), and unaccented LH (e.g. /mizu ga (H)/ ‘water + NOM’). Likewise, there are four pitch patterns in trimoraic words: HLL (e.g. /ma¹ku¹ra/ ‘pillow’), LHL (e.g. /tama¹go/ ‘egg’), accented LHH (e.g. /atama¹ ga (L)/ ‘head + NOM’), and unaccented LHH (e.g. /sakana ga (H)/ ‘fish + NOM’). For tetramoraic words, there are five patterns: HLLL (e.g. /ka¹ma¹kiri/ ‘mantis’), LHHL (e.g. /tsumeki¹ri/ ‘nail cutter’), LHLL (e.g. /kuda¹mono/ ‘fruit’), accented LHHH (e.g. /hac¹ci¹gatsu¹ ga (L)/ ‘August + NOM’), and unaccented LHHH (e.g. /tomoda¹cci ga (H)/ ‘friend + NOM’). These above examples all consist of light syllables.

There are also words containing heavy syllables. Syllables are considered heavy when they have a long vowel, diphthong, coda nasal, or geminate consonant (e.g. *yuu.ki* ‘bravery’, *ai.sa.tsu* ‘greeting’, *tom.bo* ‘dragon fly’, and *kit.te* ‘stamp’). Vance (2008) reported in his study that there was a steep pitch drop within an accented heavy syllable, whereas pitch went up within an unaccented heavy syllable or maintained a flat pitch at a high level throughout a word.

1.2 Mandarin speakers’ problems with Japanese pitch accent

It is advantageous to conduct a comparative study between L1 and L2 for a better understanding of errors related to L1 transfer. Pan (2010) investigated the characteristics of tone contours in Mandarin and made a prediction of the possible difficulties Mandarin speakers might encounter in the acquisition of Japanese pitch accent based on her findings about Mandarin contours. Mandarin has four tones (Duanmu 2002, Pan 2010): Tone 1 – high flat (pitch level 55); Tone 2 – rising (pitch level 35); Tone 3 – dip rise (pitch level 214); and Tone 4 – falling (pitch level 51). Pitch change occurs within a syllable in Mandarin unlike Japanese in which pitch change occurs within a word. Pan counted the frequency of tones occurring in 2-, 3-, and 4-syllable Mandarin words. She found that the most frequent tone pattern is falling in those words; that is, many Mandarin words start with a high pitch and end with a low pitch. Accordingly, she predicted that falling (H-L) pattern in Japanese should be easy for Mandarin speakers to pronounce, but rising (L-H) pattern should be difficult for them. Pan also pointed out that pitch drop (Tone 4) occurs within a syllable in Mandarin, which makes Mandarin speakers’ pitch span short. Considering this, Pan predicted that when pronouncing a Japanese word containing more than two syllables, Mandarin speakers might have difficulty in keeping a high pitch and may not know exactly where they

should drop pitch in a word. Likewise, Hirano et al. (2006) compared the F0 patterns of ten Tokyo Japanese and ten Mandarin speakers' readings of short Japanese sentences, and they found abrupt pitch change being made in each phrase by Mandarin speakers. In particular, Mandarin speakers of Japanese put a strong accent, consisting of a steep pitch drop, on diphthongs, which seems to be a relevant characteristic of Pan's predictions about errors.

1.3 Goal of this study

It is beneficial for teachers and learners of Japanese to obtain a good understanding of pitch patterns used in natural pronunciation. It is also important for them to be aware of common pitch errors that occur in non-native pronunciation. Being conscious of pitch contours in L2 speech will help learners improve their pronunciation and will also help educators of Japanese develop effective pedagogy and materials for teaching pronunciation.

In this study, speech data were collected from native speakers of Japanese and from Taiwan Mandarin speakers. In order to provide more extensive information about pitch contours, F0 changes on words in isolation and in carrier sentences were analyzed. For the analysis, the following questions were addressed: 1) What are the characteristics of F0 patterns in native speech? Are there any differences in pitch pattern based on environment such as a word in isolation versus a word in a sentence? 2) What types of pitch errors are observed in the speech of Mandarin speakers? Are there any commonly seen error patterns in their pitch? As Pan (2010) predicted, is falling pitch easier for them to produce than rising pitch?

2. Recording

2.1 Materials

Since the number of words with more than four moras is small in Japanese, we focused on words of 2- to 4-mora in length. We chose nouns consisting of only light syllables (e.g. /ame/ 'rain' and /tamago/ 'egg') and nouns containing at least one heavy syllable with a coda nasal, long vowel, or diphthong (e.g. /hon/ 'book' and /hikoHki/ 'airplane') for each pitch contour. As a result, twenty-seven Japanese words were selected for the recording. Table 1 shows the pitch patterns of the target words as indicated in the *Nihongo hatsuon akusent jiten* "Japanese Accent Dictionary" (Nihon, Hosoi & Kyokai 1985). The bold faced words in the table indicate the words that the Taiwanese subjects had already learned in class as taught by the author.

Table 1. List of target words

# of mora	Pitch accent pattern	Syllable structure	Words
2	HL	light – light heavy (N)	<i>ame</i> /a ^l me/ ‘rain’ <i>hon</i> /ho ^l n/ ‘book’
	LH (accented)	light – light	<i>yama</i> /jama/ ‘mountain’
	LH (unaccented)	light – light heavy (N)	<i>mizu</i> /mizu/ ‘water’ <i>ten</i> /te ⁿ / ‘point’
3	HLL	light – light – light heavy (N) – light heavy (LV) – light	<i>makura</i> /ma ^l ku ^{ra} / ‘pillow’ <i>kinko</i> /ki ^l ŋko/ ‘safebox’ <i>yuuki</i> /ju ^l ŋki/ ‘bravery’
	LHL	light – light – light light – heavy (N)	<i>tamago</i> /ta ^m ago/ ‘egg’ <i>nihon</i> /ni ^h on/ ‘Japan’
	LHH (accented)	light – light – light heavy (LV) – light	<i>atama</i> /a ^t ama/ ‘head’ <i>toori</i> /to ^h ri/ ‘street’
	LHH (unaccented)	light – light – light heavy (LV) – light heavy (N) – light	<i>sakana</i> /sa ^k ana/ ‘fish’ <i>koori</i> /ko ^h ri/ ‘ice’ <i>denwa</i> /de ^u wa/ ‘phone’
4	HLLL	light – light – light – light heavy (D) – light – light heavy (N) – light – light	<i>kamakiri</i> /ka ^l ma ^k i ^{ri} / ‘mantis’ <i>aisatsu</i> /a ^l i ^s atsu/ ‘greeting’ <i>sangatsu</i> /sa ^l ŋgatsu/ ‘March’
	LHHL	light – light – heavy (D) light – light – light – light heavy (N) – heavy (LV)	<i>kichigai</i> /ki ^c ci ^{ga} i/ ‘madness’ <i>*tsumekiri</i> /tsu ^m eki ^l ri/ ‘nail cutter’ <i>sensee</i> /se ^u se ^l / ‘teacher’
	LHLL	light – light – light – light light – heavy (LV) – light	<i>kudamono</i> /ku ^d a ^l mo ^{no} / ‘fruits’ <i>hikooki</i> /hi ^k o ^l ki/ ‘airplane’
	LHHH (accented)	light – light – light – light	<i>hachigatsu</i> /ha ^c ci ^g atsu/ ‘August’
	LHHH (unaccented)	light – light – light – light heavy (N) – light – light heavy (N) – heavy (LV)	<i>tomodachi</i> /to ^m o ^d ac ^{ci} / ‘friend’ <i>empitsu</i> /em ^p i ^t su/ ‘pencil’ <i>benkyoo</i> /be ^ŋ kjo ^h / ‘study’

Note 1: The word **tsumekiri* has another pitch pattern entry LHHH in Japanese Accent Dictionary.
2: (N), (LV), and (D) denote coda nasal, long vowel, and diphthong, respectively.

2.2 Participants

Nine university students (one male and eight females) were recruited in Taipei. Their first language was Taiwan Mandarin, and they were not fluent in Taiwan Southern Min. All of them were English majors and took a beginning Japanese course

taught by the author at National Taiwan Normal University (NTNU). Two of the subjects (subjects 4F and 7F in Table 2) had studied Japanese in high school, and the rest had been learning it only at NTNU for two semesters (about 60 hours total). Besides the Taiwanese participants, three male and three female native speakers of the Tokyo dialect, who were living in Taipei, participated in the recording. Their average length of residence in Taiwan was 5.3 months at the time of the study. Also, the author, who is a native speaker of the Tokyo dialect, recorded her speech. None of the subjects reported speech or hearing disorders. Table 2 provides relevant information about each participant.

Table 2. List of participants (F = female, M = male)

	Subject	Hometown	Length of time studying Japanese
Taiwanese	1 F	Taipei	8 months at university
	2 M	Kaohsiung	8 months at university
	3 F	Taipei	8 months at university
	4 F	Yunlin	3 years at high school and university
	5 F	Taipei	8 months at university
	6 F	Kaohsiung	1 year at university
	7 F	Yunlin	2 years at high school and university
	8 F	Taipei	8 months at university
	9 F	Taichung	8 months at university
	Subject	Hometown	Length of residence in Taiwan
Japanese	11 F	Kanagawa	6 months
	12 M	Yokohama	7 months
	13 F	Tokyo	2 months
	14 F	Tokyo	3 months
	15 M	Tokyo	5 months
	16 M	Tokyo	9 months
	17 F (author)	Chiba	4 years

2.3 Procedure

Each participant was individually recorded in a sound-attenuated booth in the phonetics lab at NTNU. Participants were seated in front of a microphone connected to a recording device.

Japanese participants were asked to look through a written word list before the recording session to ensure there were no ambiguous items, and that they were able to

pronounce all items. They were asked to read four practice words (*sora* ‘sky’, *usagi* ‘rabbit’, *shukudai* ‘homework’, and *jisho* ‘dictionary’) in order to adjust the voice volume before reading the list. Then, they read the list of the target words out loud at a comfortable speed. They put a pause of natural length (about 1 second) between items. After completing the list for the first round, they read through the list aloud once again for the second round. In the next phase, they read the same words in a carrier sentence, *kore wa ~ to yomimasu* ‘This is read as ~’ twice. All of their speech was recorded.

Taiwanese participants were given the same word list written in both Japanese letters and the Roman alphabet. They looked through the list to make sure they were able to pronounce the words. They also listened once to model pronunciations of all the words produced by a female native speaker of the Tokyo dialect (17F, the author). As soon as they finished listening to the model pronunciations, they read the words out loud for the first and second rounds. After that, they again listened to model pronunciations of all the words in the carrier sentence produced by the same native speaker, and then the subjects read the sentences aloud twice. During the recording, the native speakers of Japanese listened to their pronunciations to check for segmental mistakes. When a word was pronounced with a wrong phoneme, the subject was asked to pronounce it again.

The tokens from all the subjects were recorded at a 44.1 KHz, 16-bit sampling rate. Oral instructions were given by a female native speaker to the participants in each group in their native language. Upon the completion of the experiment, participants were asked to fill out a questionnaire about their own and their parents’ linguistic backgrounds.

2.4 Pitch measurement

Out of 1728 total tokens ((27 words × 4 repetitions) × 17 participants) recorded, only 864 tokens produced mostly in the first round were analysed. Tokens produced in the second round were used only when the production data from the first round were not usable due to mispronunciation. The F0s of vowels and coda nasals were measured at the middle points using PRAAT software (Boersma & Weenink 2014). The F0s of long vowels were measured at the middle and offset points. Changes in pitch levels were calculated by subtracting the F0 of the preceding mora from the F0 of the following mora. Also, the F0 of the first mora was subtracted from the F0 of the last mora in the 3- and 4-mora words.

3. Results

Pitch patterns of the tokens were analyzed in order to 1) verify how the actual

pitch patterns in native speech are compatible or incompatible with the prevailing notions of Japanese pitch accent in language textbooks; and to 2) detect how the learners preserve or distort the Japanese pitch accent patterns. As a broad overview of the analysis, the native speakers did not always exhibit the same pitch pattern in the words in isolation and words used in sentences. It is reasonable to adjust the pitch patterns of the words in sentences in order to produce natural intonation as a whole. As for the Taiwanese subjects, they tended to make pitch errors more frequently in sentences than in the isolated words. Since their Japanese proficiency was at the beginner level, the longer the utterance, the more difficulty they had in controlling pitch in the right way. A more detailed analysis of the pitch patterns of the target words is provided below.

3.1 Pitch patterns of two-mora words

In Figure 1, the line graphs display the pitch levels in each mora of the HL words (*ame* and *hon*) sorted according to the subject group (Taiwanese and Japanese) and the presentation of the target word (in isolation and in a sentence). As we can see, all the learners had the right pitch accent for the word *hon*, but some produced a totally opposite rising pitch (LH) for the word *ame*. This may be related to the familiarity of the words. Since the word *hon* was frequently used in class, the learners had already mastered its accent, but they had not mastered the accent of *ame* perhaps due to the lack of enough exposure although they heard a model pronunciation right before the recording.

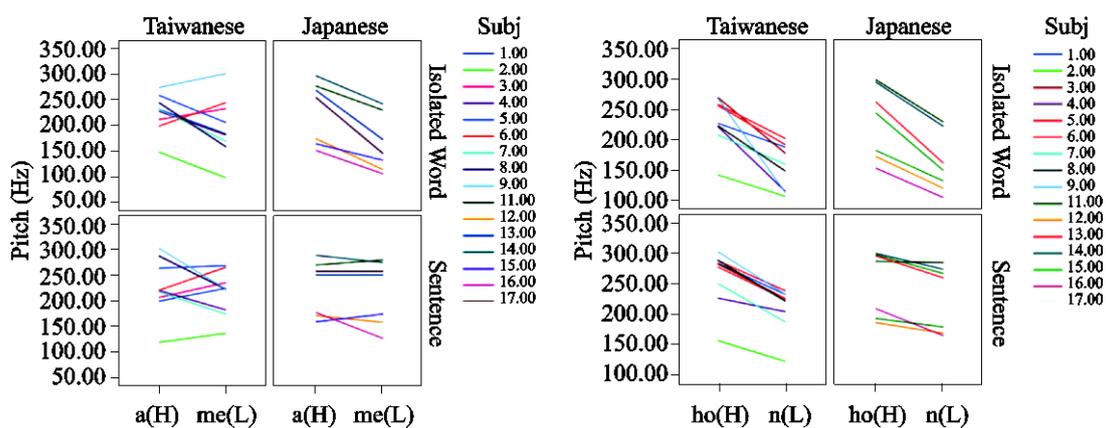


Figure 1. HL pitch pattern for *ame* (left) and *hon* (right) for each subject

A look at the data from the native speakers shows that the pitch drop from the first to the second mora in HL was notably not very steep in a sentence (-15 Hz and -24.3 Hz

on average for the words *ame* and *hon*, respectively) compared with the words in isolation (-63.4 Hz and -69.3 Hz, respectively). Instead of lowering the second mora, the native speakers steeply lowered pitch on the following syllable *to* ‘quoting particle’ in a sentence, which can mark the border of an intonational phrase (IP). The learners, on the other hand, tended to make a steep pitch drop within the target word in a sentence (-56.5 Hz for *ame* and -51.8 Hz for *hon*) and keep the same pitch level on the following particle *to*. The up spectrogram of Figure 2 displays an utterance of *kore wa hon to | yomimasu* ‘This is read as book’ by a native speaker (17F). The first IP consists of an unaccented accentual phrase (AP) *kore wa* and an accented AP *hon to*. The pitch of the target word greatly rose at the beginning of the accented AP and steeply dropped in the transition from the end of the second mora [n] to the next syllable, which shows a delayed F0 fall (*oso-sagari*) and marks the border of an IP. By contrast, a learner (see the down spectrogram of Figure 2) gradually raised pitch from the preceding particle *wa* ‘topic particle’ to the first mora of the target word, dropped pitch on the second mora, and maintained flat pitch through the following particle. This pattern obscures the borders between the adjacent particles and target word, and hence IP. Even though the learner knew the right pitch accent of the target word, she did not adjust pitch in a sentence to form a clear IP like the native speaker did.

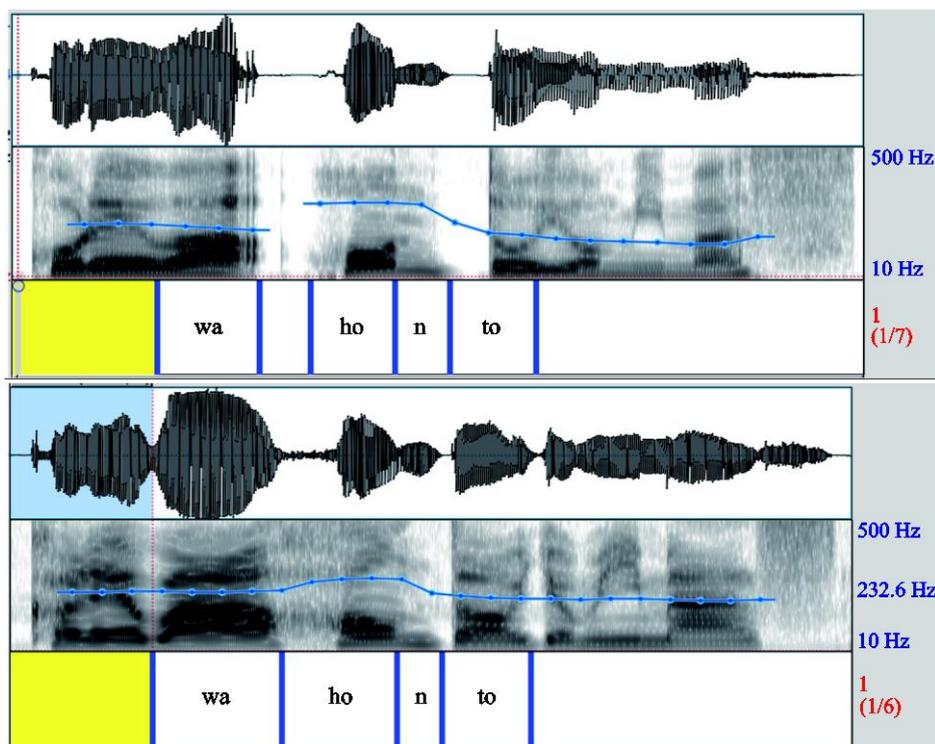


Figure 2. Up: Pitch of *hon* by a native speaker (17F), Down: Pitch of *hon* by a learner (5F)

Figure 3 shows pitch patterns of the LH words (*yama*, *mizu*, and *ten*). There were two error patterns observed in non-native speech: making a sharp fall (HL) or a low flat pitch without a clear pitch rise. Actually, pronunciation of the word *ten* with a low flat pitch was observed in native speech as well. In the Japanese data, the pitch rise from the first to the second mora in the accented LH word *yama* was almost twice as high in a sentence (+64.7 Hz) compared to the isolated word (+33.4 Hz). Since the second mora *ma* must carry an accent, a steep rise makes a more distinct contrast with the following low-pitched mora *to* in a sentence. It adds more prominence to the accented mora. On the other hand, in the unaccented LH words (*mizu* and *ten*), the amount of pitch rise was not very different between the target word in a sentence and the word in isolation. For *mizu*, the pitch rise in the isolated word and in a sentence was +40 Hz and +35.1 Hz, respectively, and for *ten*, the pitch rise in the isolated word and in a sentence was only +9.5 Hz and +10.3 Hz, respectively. As mentioned above, pitch was rather low flat for the word *ten*, which consists of a single heavy syllable.

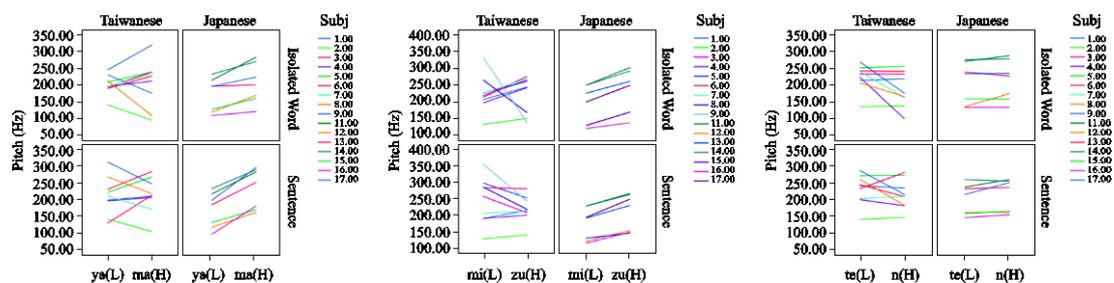


Figure 3. Accented LH pitch pattern for *yama* (left), unaccented LH pitch pattern *mizu* (middle) and *ten* (right)

3.2 Pitch patterns of three-mora words

An error commonly observed in the HLL words (*makura*, *kinko*, and *yuuki*) among the learners was starting the words with low pitch followed by raised pitch, making either LHL or accented LHH (see Figure 4). Among the three words, the word *makura* was pronounced incorrectly most of the time (five subjects for the isolated word and six for the word in a sentence) compared with the other two words containing a heavy syllable. It seems that dropping pitch at one stretch within a heavy syllable like Mandarin Tone 4 was easier for the learners. Most of the tokens with falling pitch exhibited a sharp pitch drop all the way through the words (see Figure 5).

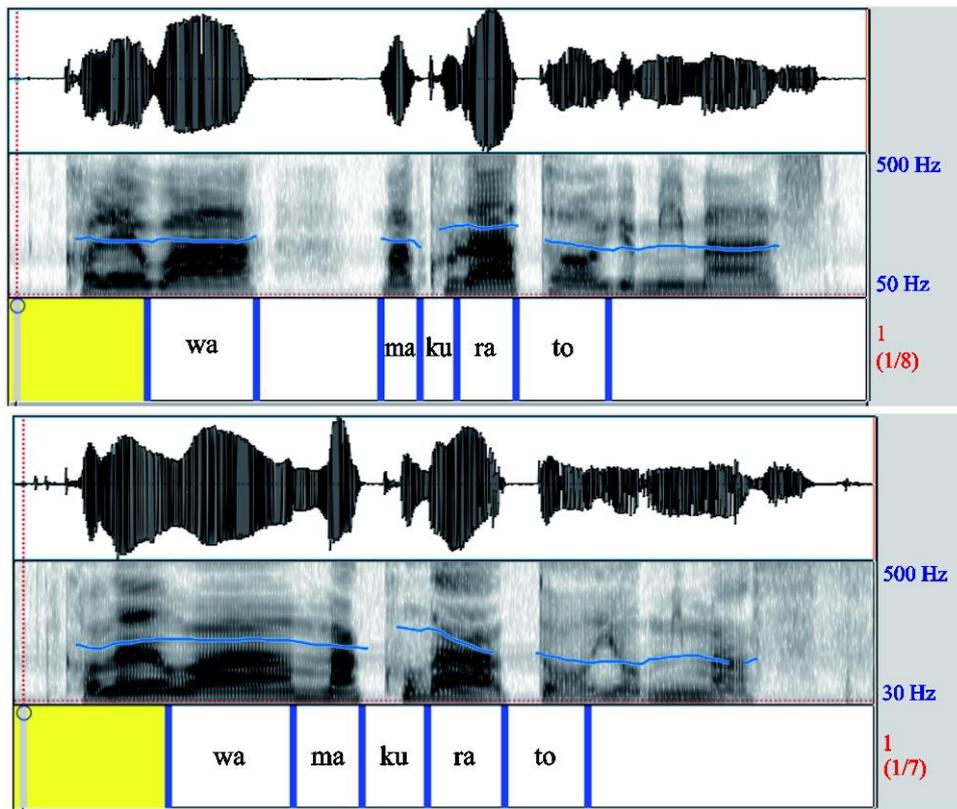
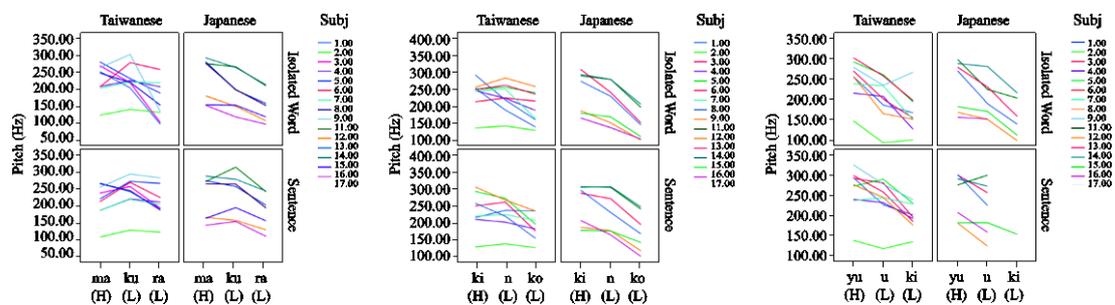


Figure 4. Up: *makura* (incorrect LHH) by 5F, Down: *makura* (incorrect LHL) by 6F



Note: A mora with no line denotes no measured pitch due to vowel devoicing

Figure 5. HLL pitch pattern for *makura* (left), *kinko* (middle), and *yuuki* (right)

The HLL words produced by the native speakers revealed two patterns of pitch drop: 1) a steep pitch drop all the way through the words; or 2) a gradual pitch drop from the first to the second mora and a sharp drop from the second to the third mora, which looks like HHL by accent shift. More interestingly, three native speakers even raised pitch on the second mora in the word *makura*, such that the pitch pattern became LHL in a sentence, and again the accent shifted to the second mora.

Figure 6 shows pitch patterns for the LHL words (*tamago* and *nihon*). The native speakers dropped pitch drastically on the third mora (the average -74.9 Hz in *tamago* and -75.4 Hz in *nihon*) for the word in isolation, which resulted in a Mid-High-Low pattern. On the other hand, pitch was raised steeply on the second mora (the average $+54.7$ Hz in *tamago* and $+73.9$ Hz in *nihon*) in the word in a sentence and pitch did not drop much on the third mora, producing a Low-High-Mid pattern. Adjusting pitch to mid level on the last mora leaves some room for a further pitch drop on the following mora, which helps make the IP border distinct.

All the learners made the correct accent for *nihon*, which was a very familiar word to them. Similar to the native speakers' pronunciations, the pitch pattern of the learners was Mid-High-Low for the word in isolation and Low-High-Mid for the word in a sentence. The only slight difference was that pitch rise on the second mora (average $+52.1$ Hz) was less than that in the native speakers' pronunciation ($+73.9$ Hz). The other LHL word, *tamago*, received more error patterns such as HLL, HHL, or a low flat pitch all the way through the word.

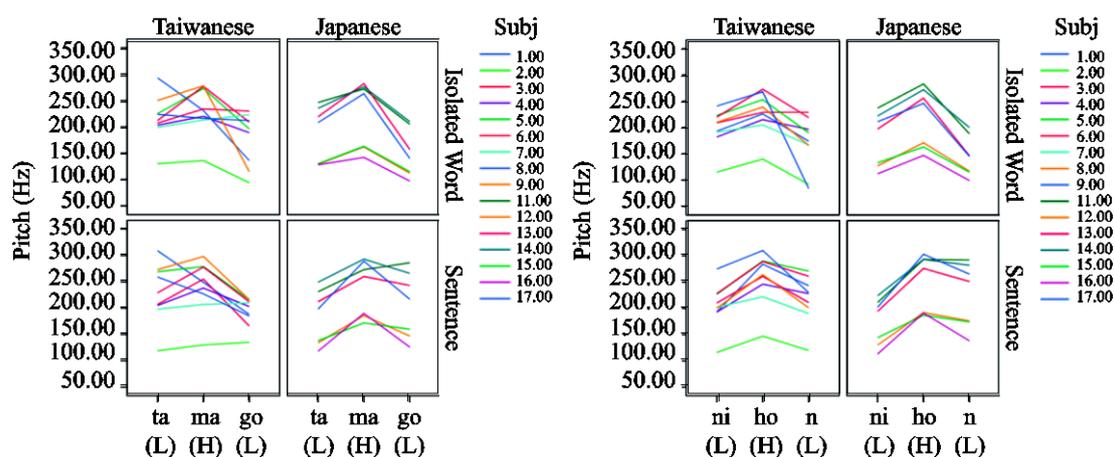


Figure 6. LHL pitch pattern for *tamago* (left) and *nihon* (right)

For the accented LHH word *atama* (see Figure 7), the native speakers raised pitch from the first to the second mora to a greater degree for the word in a sentence (average $+53.1$ Hz) than for the isolated word ($+39.7$ Hz). After a rise on the second mora, pitch level on the third mora seems to vary among the speakers: some kept pitch rising, some kept pitch flat, and some lowered pitch, making LHL. In a sentence, only one subject clearly made the LHL pattern, and the rest raised pitch or kept it flat on the third mora. Compared with the native speakers' pronunciation, the pitch rise on the second mora by the learners was less clear, especially in the isolated word. Moreover, some learners produced an LHL pattern. More

importantly, two learners produced the correct pattern for the target word but did not lower pitch on the following particle, resulting in an unaccented LHH pattern in a sentence.

For another accented LHH word *toori* containing a long vowel, the native speakers did not make a clear LHH pattern for this word as shown in Figure 7. Pitch did not rise within the long vowel, and it stayed at low flat or even lower on the third mora for most of the native speakers. This tendency was more obvious for the word in isolation than the word in a sentence. Some learners also kept a low flat pitch all the way through this word while others even made a steep drop on the second mora, producing HLL.

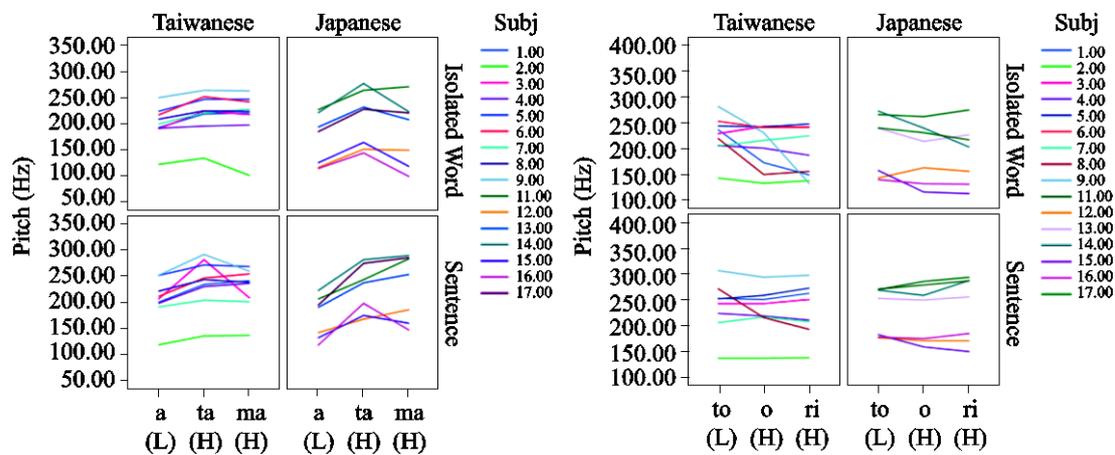


Figure 7. Accented LHH pitch pattern for *atama* (left) and *toori* (right)

Figure 8 displays the pitch patterns of native speakers in the unaccented LHH words (*sakana* and *denwa*). Like the accented LHH word *atama*, pitch rise by the native speakers was greater (+32.1 Hz for *sakana* and +30 Hz for *denwa*) for the word in a sentence than on the isolated word (+26.6 Hz for *sakana* and +10.6 Hz for *denwa*), and the pitch level on the third mora was inconsistent among the speakers. For some native speakers, pitch kept rising; some lowered pitch, and some kept a flat pitch. It is clear in the figures that the pitch patterns for the other unaccented LHH word *koori* (Figure 8) were very similar to the accented LHH word *toori* (Figure 7). Pitch was not raised on the second mora and either a low flat pitch all the way through the word or a subtle pitch rise on the third mora was observed. As Figure 9 shows, among all the three words containing a long vowel in the first syllable (*yuuki*, *toori*, and *koori*), only falling pitch exhibited the expected pattern (HLL); in actuality, LHH did not produce the rising contour expected of native speaker pronunciation.

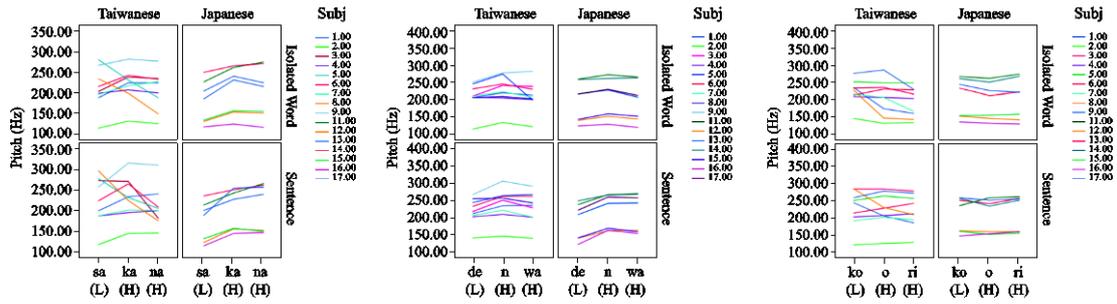


Figure 8. Unaccented LHH pitch pattern for *sakana* (left), *denwa* (middle), and *koori* (right)

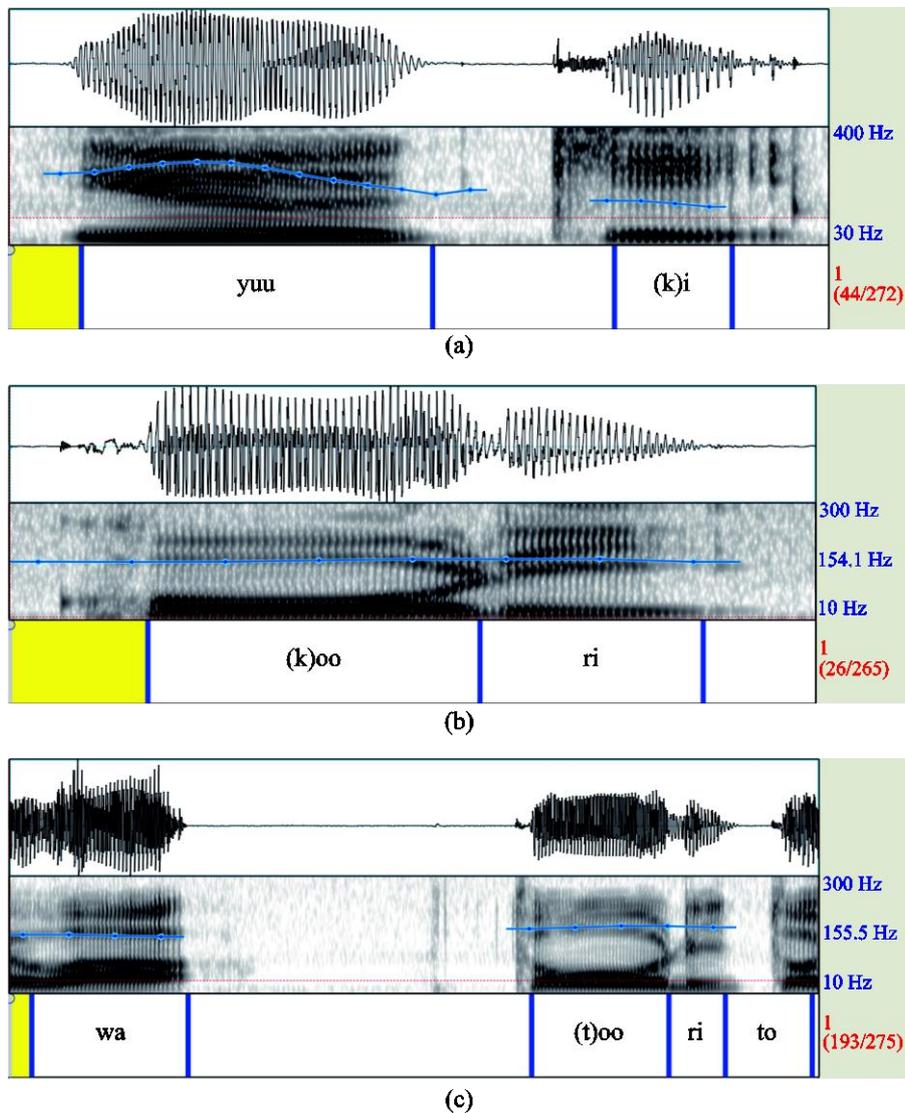


Figure 9. (a): HLL (*yuuki*) in isolation produced by 13F, (b): unaccented LHH (*koori*) in isolation produced by 15M, (c): accented LHH (*toori*) in a sentence produced by 12M

Back to Figure 8, common pitch errors made by the learners for the unaccented LHH words (*sakana*, *denwa*, and *koori*) were either putting an accent on the second mora (LHL) or putting an accent on the third mora by lowering pitch after the target word in a sentence, making an accented LHH pattern (see Figure 10). It seems that the learners were not aware of the distinction between accented and unaccented LHH pitch. Other than that, mistakes appeared to be random; some subjects produced HLL, LHL, or HHL patterns. As mentioned above, rising pitch within a long vowel tends to be low flat without a significant rise in the native speakers' speech, so these words (*toori* and *koori*) might be easier for the learners to pronounce native-like.

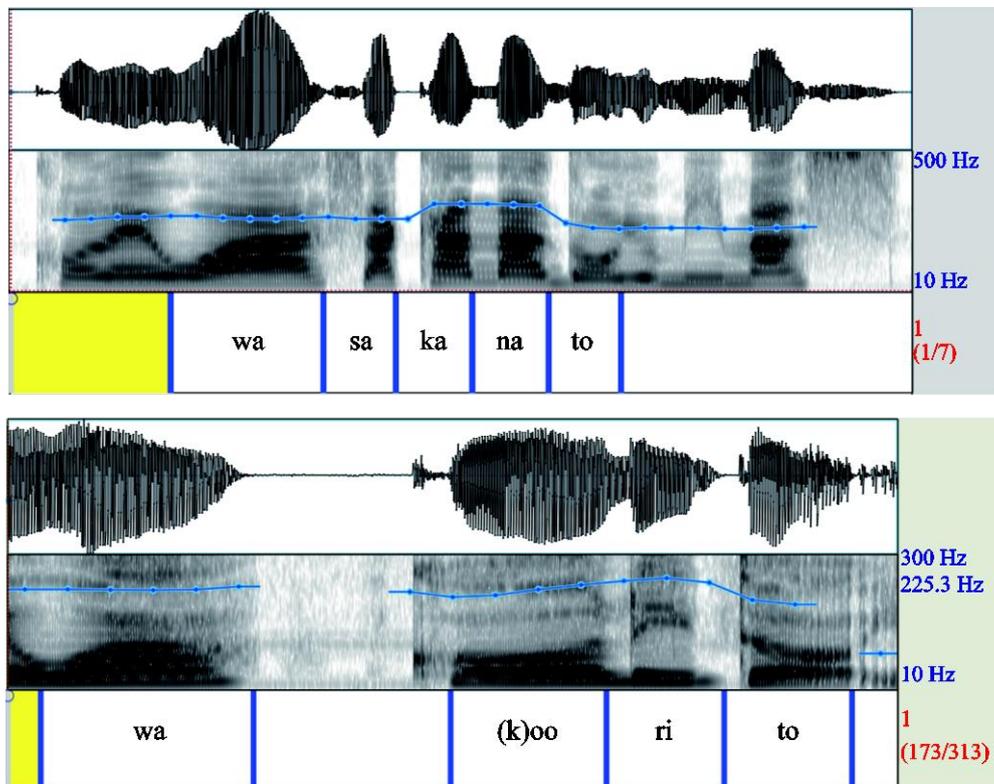


Figure 10. Up: Accented LHH *sakana* produced by 9F, Down: Accented LHH *koori* produced by 3F

3.3 Pitch patterns of four-mora words

Similar to the three-mora words with falling pitch, the HLLL words (*kamakiri*, *sangatsu*, and *aisatsu*) produced by the native speakers show two patterns of pitch drop: 1) a steep pitch drop all the way through the words; or 2) a moderate pitch drop from the first to the second mora, a sharp drop from the second to the third mora, and

a moderate drop from the third to the last mora, which demonstrates that the pattern was actually HLL including accent shift (see Figure 11). All the native speakers had the lowest pitch on the third mora of these words when in a sentence, and more than half of them made a flat pitch from the first to the second mora. Only the word *aisatsu*, containing a diphthong in the first syllable, behaved differently in the isolated word; six out of seven speakers lowered pitch the most on the second mora instead of the third mora.

Unlike the native speakers, all the learners raised pitch on the second mora of the word *kamakiri* and produced the lowest pitch on the last mora (LHHL) as shown in Figure 11. Similar characteristics were observed for the words *sangatsu* and *aisatsu* containing a heavy syllable. Six to seven non-native speakers either raised pitch or made a flat pitch from the first to the second mora and dropped pitch on the third or the last mora.

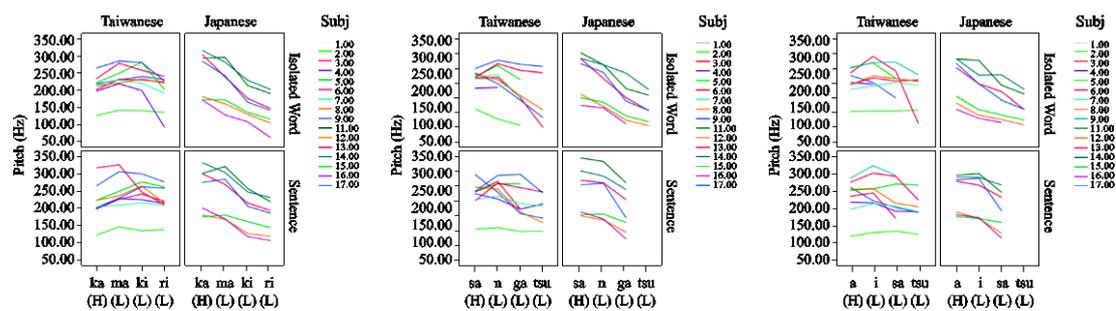
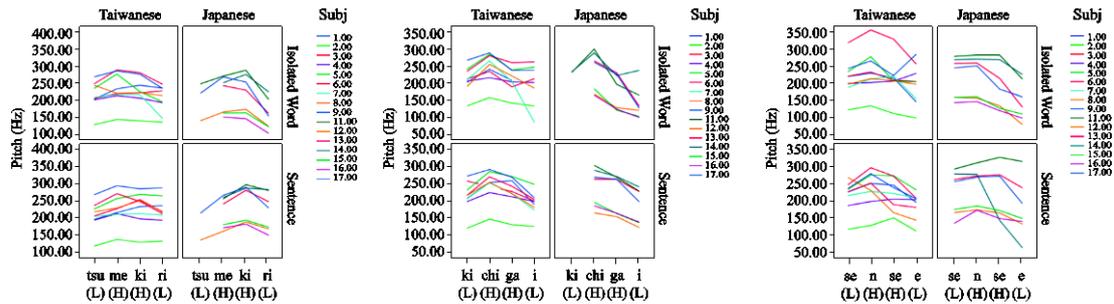


Figure 11. HLLL pitch pattern for *kamakiri* (left), *sangatsu* (middle), and *aisatsu* (right)

Figure 12 shows data for the LHHL words (*tsumekiri*, *kichigai*, and *sensee*). All the Japanese speakers dropped pitch on the last mora of *tsumekiri*, making a clear LHHL pattern. Pitch drop was greater for the isolated words (average -67.3 Hz) than for the words in a sentence (-26.6 Hz). The learners were able to raise pitch on the second mora, but many of them either kept a flat pitch until the end of the word (LHHH) or dropped pitch on the third mora (LHLL). Although pitch drop on the third mora by the learners was observed for the other two words (*kichigai* and *sensee*), it did not actually sound like an error because the native speakers had also made the same LHLL contour. Many native speakers dropped pitch on the first part of a heavy syllable, so the position of the pitch drop shifted to the third mora. Also, pitch within the heavy syllable in *sensee* had a slight rise or was low and flat for the isolated word.

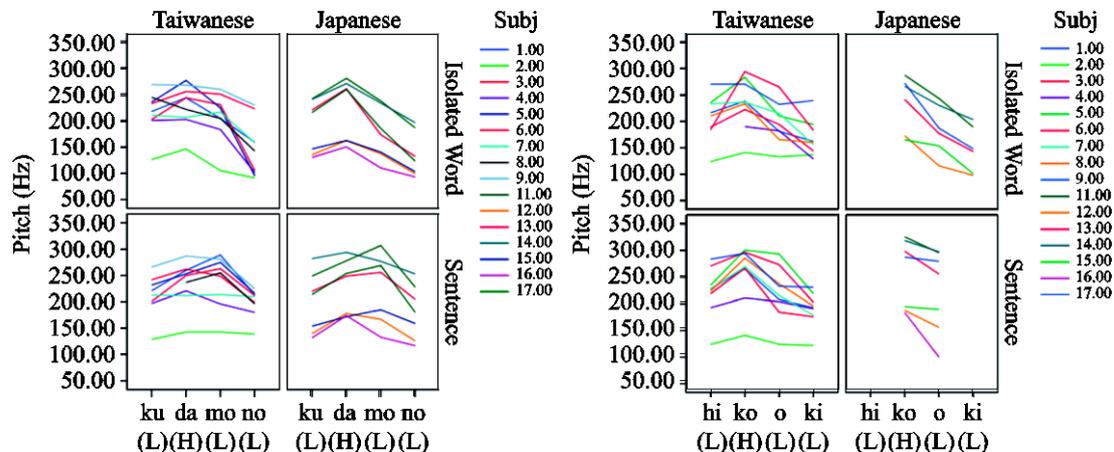


Note: Mora with no pitch line denotes vowel devoicing

Figure 12. LHL pitch pattern for *tsumekiri* (left), *kichigai* (middle), and *sensee* (right)

More than half of the native speakers produced the LHL pitch pattern for the word *kudamono* with accent shift in a sentence (see Figure 13). They dropped pitch on the last mora instead of the third one (LHL) while all of them had the LHL pattern for the isolated word. Another word, *hikooki*, lacked pitch data for the first and the last moras due to vowel devoicing, but pitch drop within the long vowel was visible. Indeed, the word *hikooki* seems easier than the word *kudamono* for the learners to correctly pronounce. Almost all of the learners had the correct pitch pattern for *hikooki*, except that some tokens sounded rather flat from the third to the fourth mora, whereas most of them pronounced *kudamono* incorrectly with either high flat LHHH, or accented LHL (see Figure 14).

Lastly, Figure 15 displays pitch patterns of the accented LHHH word (*hachigatsu*) and unaccented LHHH words (*tomodachi*, *empitsu*, and *benkyoo*). The Japanese speakers raised pitch from the first to the second mora, and the accented LHHH word *hachigatu* has the greatest pitch rise: an average +55.1 Hz for the isolated word and



Note: Mora with no pitch line denotes vowel devoicing

Figure 13. LHL pitch pattern for *kudamono* (left) and *hikooki* (right)

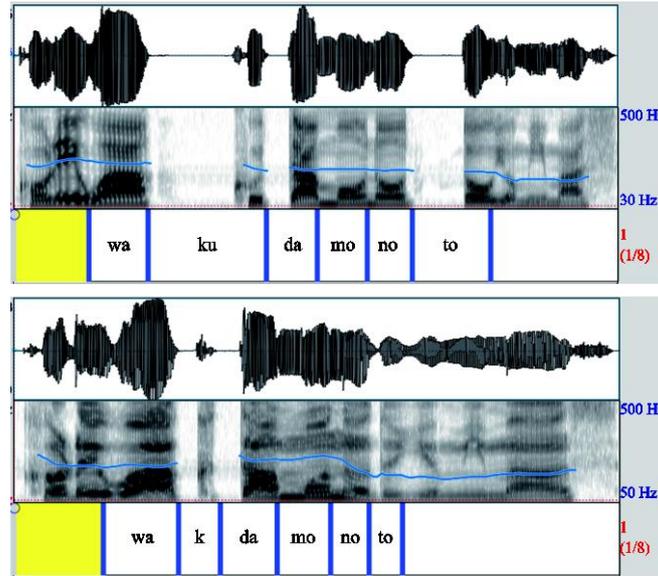
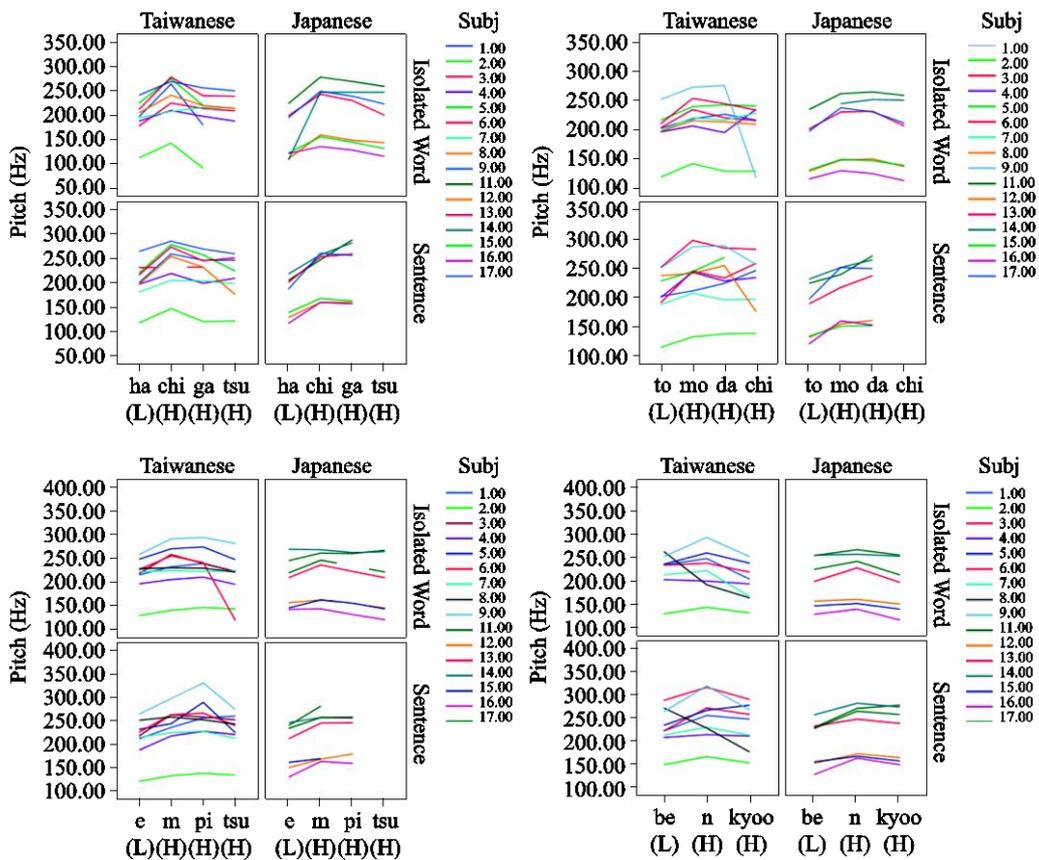


Figure 14. Up: High flat *kudamono* produced by 7F, Down: LHHL (L) *kudamono* (to) produced by 8F



Note: Mora with no pitch line denotes vowel devoicing

Figure 15. Accented LHHH *hachigatsu* (top left) and unaccented LHHH (H) *tomodachi* (top right), Unaccented LHHH (H) *empitsu* (bottom left) and *benkyoo* (bottom right)

+45.6 Hz for the word in a sentence. Compare this with the unaccented LHHH words (e.g. *tomodachi*, +24.5 Hz for the isolated word and +28 Hz for the word in a sentence; *empitsu*, +13.3 Hz for the isolated word and +23.9 Hz for the word in a sentence; and *benkyoo*, +11.6 Hz for the isolated word and +26.4 Hz for the word in a sentence). There was a flat pitch or slight pitch rise or fall observed from the third to the fourth mora. This was a tendency commonly seen, the exception being the word *benkyoo* in isolation. The word *benkyoo* showed a different pitch pattern: pitch dropped an average -17.3 Hz on the third mora (the first half of a long vowel), making the LHLL pattern.

For the words *hachigatsu*, *tomodachi*, and *empitsu*, there were a few distorted patterns observed among the learners. Some of them lowered pitch drastically on the third or the fourth mora, producing a LHLL or LHHL pattern; or a few learners produced an LHLH pattern, which is an impossible accent pattern in Japanese. Moreover, since the learners seemed unaware of the distinction between accented and unaccented rising pitch, pitch on the following particle *to* after the target word was not well-controlled. Some did not lower pitch after the accented word whereas others lowered pitch after the unaccented word as seen in Figure 16.

4. Summary and Discussion

One of the goals of this study was to examine the characteristics of Japanese pitch contours in L1 speech. The analysis of F0 patterns in 2-, 3-, and 4-mora words yielded several findings about the production of native speakers of Japanese.

First, in the falling pitch contours (HL, HLL, and HLLL), the range of pitch drop is greater for words in isolation than for the same word in a sentence. This is not surprising because pitch usually tends to drop toward the end of an utterance (Herman 2000). In the sentence pattern, an unaccented AP *kore wa* ‘this is’ and the following AP (a target word followed by the quoting particle *to*) form an IP. Pitch drop on the last mora of the target noun is less steep so that there is still room for a further pitch drop on the following particle to mark an IP boundary.

Second, two patterns of pitch drop are found in longer words. Pitch falls in a stepwise fashion all the way through the word, or pitch drops moderately from the first to the second mora followed by a steep drop from the second to the third mora. The latter pattern can be attributed to delayed F0 fall, causing a HHL or HHLL pattern with accent shift. Although having the former or latter pattern seems to be ideolectal variation, the latter case with a delayed F0 fall occurred more commonly and frequently for words in sentences.

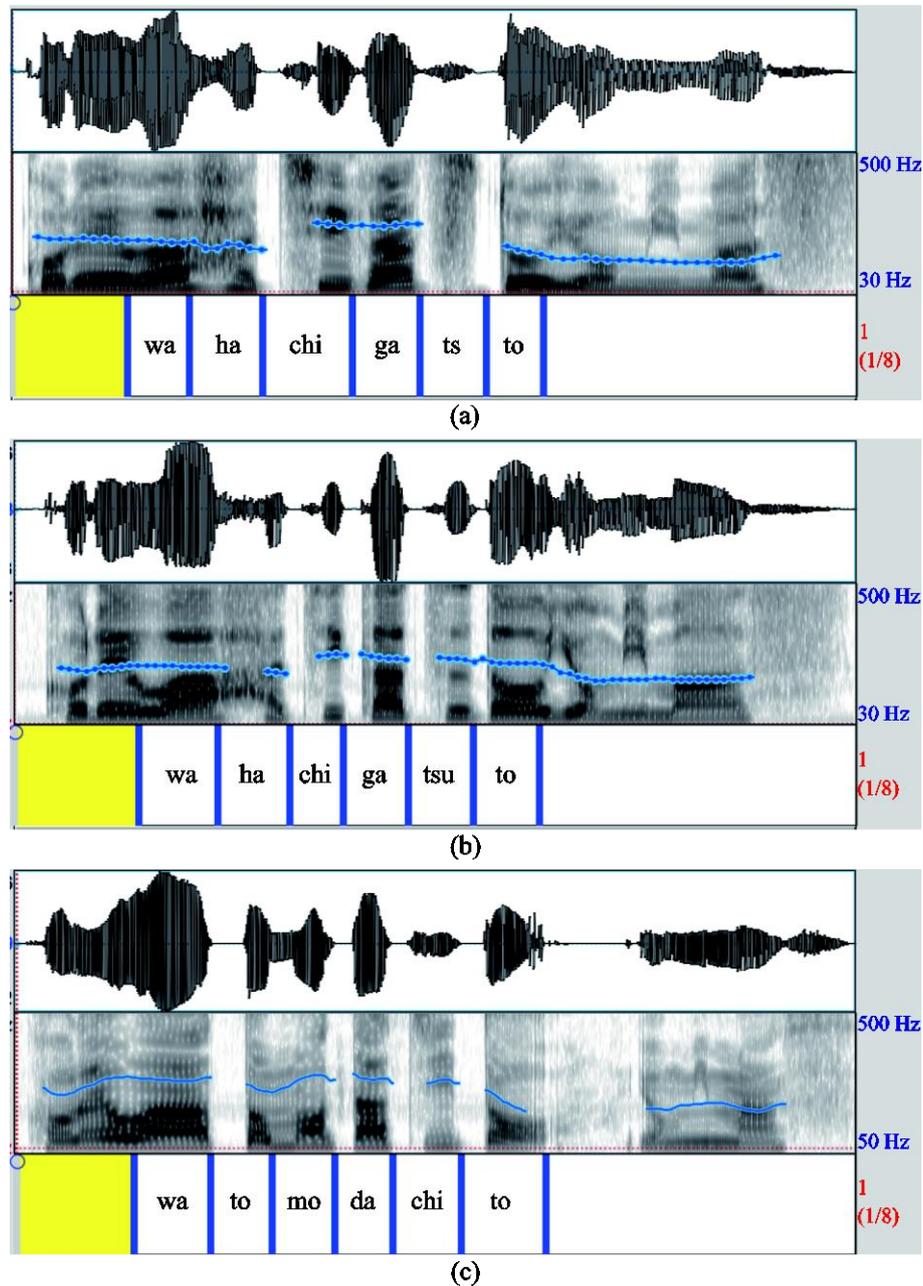


Figure 16. (a): Accented *hachigatsu* by a Japanese (17F), (b): Unaccented *hachigatsu* by a learner (1F), (c): Accented *tomodachi* by a learner (6F)

Third, rising pitch contours showed a much steeper rise at the beginning of the words in a sentence than those same words in isolation, this being more obvious for accented words than for unaccented words. Pitch level after the second mora is random; rising, slightly falling, and high flat can be observed. It is interesting to note that for a falling contour, a more abrupt pitch change occurs for isolated words than for the words in sentences; conversely, for a rising contour, abrupt pitch change occurs more frequently for the words in sentences than for isolated words.

Fourth, rising-falling contours (LHL, LHHL, and LHLL) inherit the characteristics of simple rising and falling contours. There is a steeper rise from the first to the second mora for words in sentences than for words in isolation, and a steeper drop toward the end of the words for words in isolation than for words in sentences. Thus, this contour appears to be Mid-High-Low for the words in isolation and Low-High-Mid for the words in sentences.

Finally, rising pitch patterns of heavy syllables seem somewhat different from those of light syllables. Actually, there is no clear pitch rise within a heavy syllable; rather, pitch stays low flat or a slight drop is even observed. This finding is compatible with an observation made by Vance (2008).

Another goal of this study was to discern the learners' difficulties in producing Japanese pitch contours. Since they were all beginners of Japanese, they made all kinds of pitch errors, such as producing the opposite pitch contour, putting an accent in the wrong position, and having two accents on a word.

Although the errors seemed random, some common problems could still be detected. First, based on the number of errors, those words consisting of all light syllables seem more difficult to pronounce with the accurate pitch contour than the words containing a heavy syllable. In Mandarin, pitch changes at the syllable level; thus, forming one pitch contour at the word level must be more difficult for the learners as the number of syllables in a word increases.

Second, it was found that the learners tended to produce a steep pitch drop all the way through the words. As mentioned before, Pan (2010) predicted that Mandarin speakers should be good at producing falling contours in Japanese because Tone 4 is the most frequent tone at any position in a word. The finding in the current study supports her prediction. Pan also pointed out that Mandarin speakers may have difficulty in maintaining a high pitch and may not know where they should drop pitch in a longer word. This was also observed in the LHHH words in the present study. Some learners did not maintain high pitch through to the end of a word, instead, drastically dropped pitch after the second mora.

Lastly and most importantly, the learners seemed unaware of the difference between accented and accent-less rising pitch. Accented rising pitch must be followed by a low pitch, but many learners did not lower pitch on the following mora. Likewise, learners made the mistake of unnecessarily lowering pitch on the mora following the unaccented words. This suggests that even though they have good control of pitch at the word level, they are not able to adjust pitch at the phrase level.

The results from this study suggest that Japanese pitch contours are not always constant. The patterns change based on multiple factors, such as ideolectal variation, environment (words in isolation versus those in a sentence, in the case of this study),

and the number of moras in a word. This study also found that falling pitch contours are relatively easy for Mandarin speakers to pronounce. On the other hand, for rising and rising-falling contours, maintaining a high pitch and dropping pitch on the correct mora seem difficult for them. Additionally, it was revealed that they were not aware that pitch on the following mora determines whether a word is accented or unaccented.

From the above findings, the importance of learning/teaching the pitch accent system should be emphasized in the classroom. Visualizing F0 contours in a model sentence and practicing pronunciation using such visual aids should be an effective way for acquisition of pitch accent as suggested in some literature (Hirata 2004, Hirano 2014). In order for learners to successfully acquire correct pitch, it is necessary to practice pitch patterns not only at the word level but also at the phrase and sentence levels at an early stage of learning. For this, instructors should obtain a good knowledge of Japanese pitch accent by reviewing not only language textbooks but also previous research on the topic. Also, it may be a good approach to collect common errors that learners often make and create a database of such errors. The database can be shared among instructors of Japanese. The information about errors should be useful for teachers to anticipate or detect errors in their actual teaching situations. In addition, instructors should raise learners' awareness of their pitch errors by carefully listening to their speech and correcting errors in the classroom. Providing a model speech and getting learners to repeat it may be useful, too. It is also effective to get learners to read a short passage aloud and give evaluation on it so that learners are able to realize how their pronunciation is perceived from another person's point of view.

5. Conclusions

The main purpose of this study was to provide an analysis of the Japanese pitch accent system in actual L1 and L2 speech. The study discovered peculiarities of F0 contours in native speech and pointed out difficulties in the acquisition of Japanese pitch accent observed in Mandarin speakers' speech. It was suggested that learners should become aware of pitch contours at phrase and sentence levels, and they should receive sufficient pronunciation practice in order to acquire 'good' native-like pronunciation. This study merely focused on how F0 contours appear in L1 and L2 pitch contours. However, a more thorough study on how Mandarin tone features affect acquisition of Japanese pitch accent is desired. Since Taiwan Mandarin and Japanese share many words written in the same characters (e.g. 安心 'ease', 警察 'police'), it is possible that tone contours of those words in Mandarin might influence pitch contours in Japanese reading by Mandarin speakers. Future work should

continue to investigate this issue to provide more useful data for a better understanding of prosodic errors in L2 speech.

References

- Akinaga, Kazue. 1966. Kyootsuugo no akusento [Accent of standard Japanese]. *Nihongo Hatsuon Akusento Jiten [Japanese Accent Dictionary]*, ed. by NHK Hoosoo Bunka Kyookai, 174-221. Tokyo: NHK Shuppan.
- Anderson-Hsieh, Janet, Ruth Johnson, and Kenneth Koehler. 1992. The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning* 42.4:529-555.
- Boersma, Paul. and David Weenink. 2014. *Praat: Doing Phonetics by Computer, Version 5402 (computer software)*. Amsterdam: Institute of Phonetic Sciences, University of Amsterdam.
- Duanmu, San. 2002. *The Phonology of Standard Chinese*. New York: Oxford University Press.
- Hasegawa, Yoko. 1995. Against marking accent locations in Japanese textbooks. *Sekai no Gengokyoiku [World Language Education]* 5:95-103.
- Hata, Kazue and Yoko Hasegawa. 1988. Delayed pitch fall in Japanese: A perceptual experiment. *The Journal of the Acoustical Society of America*. 84.S1:S156-S156.
- Herman, Rebecca. 2000. Phonetic markers of global discourse structures in English. *Journal of Phonetics* 28.4:466-493.
- Hirano, Hiroko. 2014. Soogoo Nihongo no jyugyoo de okonau zero shokyuu kara no onseikyooiku no jissen: Akusento, intoneeshon no shizensee wo jyuusi shita shikakukahojokyoozai no shiyoo [Practice of teaching Japanese phonetics from novice level in general Japanese classes: Usage of visual aids focusing on naturalness of accent and intonation]. *National Institute for Japanese Language and Linguistics Research Papers* 7:45-71.
- Hirano, Hiroko, Keikichi Hirose, Nobuaki Minematsu, and Goh Kawai. 2006. Prosodic features and evaluation of Japanese sentences spoken by Chinese learners. *IEICE Technical Report, Speech* 105(686):23-28.
- Hirata, Yukari. 2004. Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning* 17.3-4:357-376.
- Hirayama, Teruo. 1966. Zennihon no hatsuon to akusento [Pronunciation and accent in the whole Japan]. *Nihongo Hatsuon Akusento Jiten [Japanese Accent Dictionary]*, ed. by NHK Hoosoo Bunka Kyookai, 123-173. Tokyo: NHK

Shuppan.

- Ishihara, Shunichi, Yiran Fan, Dia Jalil, and Chiharu Tsurutani. 2011. A comparative study on perception of foreign-accented Japanese by L2 Japanese listeners having different L1 backgrounds: English, Chinese, and Indonesian. *Proceedings of the 17th International Congress of Phonetic Sciences*, ed. by Wai-Sym Lee and Eric Zee, 950-953. Hong Kong: City University of Hong Kong.
- Japan Foundation. 2013. *Wochi Kochi Magazine*. Retrieved September 20, 2014 from <http://www.wochikochi.jp/topstory/2013/12/learning-japanese.php>.
- Japan Foundation. 2014. *Japanese-Language Education Overseas*. Retrieved September 20, 2014 from <http://www.jpf.go.jp/j/japanese/survey/country/2013/taiwan.html>.
- Jilka, Matthias. 2000. Testing the contribution of prosody to the perception of foreign accent. *Proceedings of New Sounds 2000*, ed. by Allan James and Jonathan Leather), 199-207. Amsterdam: University of Amsterdam.
- Kindaichi, Haruhiko. 1966. Kyootsuugo no hatsuon to akusento [Pronunciation and accent in Standard Japanese]. *Nihongo Hatsuon Akusento Jiten [Japanese Accent Dictionary]* ed. by NHK Hoosoo Bunka Kyookai, 91-122. Tokyo: NHK Shuppan.
- Kubozono, Haruo. 2008. Japanese accent. *Handbook of Japanese Linguistics*, ed. by Shigeru Miyagawa and Mamoru Saito, 165-191. Oxford: Oxford University Press.
- Kubozono, Haruo. 2012. Varieties of pitch accent systems in Japanese. *Lingua* 122.13: 1395-1414.
- Munro, Murray J., and Tracey M. Derwing. 1995. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 45.1:73-97.
- Neustupný, Jiří Václav. 1966. Is the Japanese accent a pitch accent? *Onsei Gakkai Kaihoo* 121. Reprinted in: Tokugawa, Munemasa. 1980. *Ronshu Nihongo Kenkyuu 2 Akusento [Papers on Japanese Language Research 2, Accent]*, 230-239. Tokyo: Yuuseido.
- NHK Hoosoo Bunka Kyookai (ed). 1985. *Nihongo Hatsuon Akusento Jiten* (1st edition) [*Japanese Accent Dictionary*]. Tokyo: NHK Shuppan.
- Pan, Hsin-ying. 2010. Pekingo no seichootokuchoo kara yosoku suru pekingowasha ni okeru akusentoshuutoku no mondaiten [Chinese Mandarin native speakers' accent-learning problems predicted according to characteristics of tones in Mandarin] *Gengogaku Ronsoo [Translation]* 29:18-31.
- Shport, Irina A. 2008. Acquisition of Japanese pitch accent by American learners. *Japanese as Foreign Language in the Age of Globalization*, ed. by Patrick Heinrich and Yuko Sugita, 165-187. German Institute for Japanese Studies, München: Iudicium Verlag.
- Sugito, Miyoko. 1970. Nihongoboin no dootaisokutei to akusento no ninshiki

[Measurements of tone movement of vowels and hearing validity in relation to accent in Japanese] *Studia Phonologica* 5:1-19.

Vance, Timothy J. 2008. *The Sounds of Japanese*. New York: Cambridge University Press.

[Received April 14, 2015; revised July 1, 2015; accepted October 22, 2015]

Center for Language Research
University of Aizu
Aizuwakamatsu, JAPAN
Naomi Ogasawara: naomi-o@u-aizu.ac.jp

台灣華語人士習得日語之音高重音分析

小立原奈保美

會津大學

本研究旨在探討台灣日語學習者在初期階段音高類型之習得，並比較其與日語母語人士之差異。研究目標在於觀察台灣日語學習者在字面與句面層次中的日語音高類型之特點，且進一步歸納出常見之錯誤音高類型。由研究成果可得知，台灣日語學習者較難掌握完全由輕音節組成之詞彙，也常會在字內有音高陡降的情形。更重要的是，學習者似乎無法察覺有重音之上揚音高後接低音以及無重音之上揚音高後接高音之差別。藉此可得知，即使學習者知道字面上的正確日語發音，但實際口說方面仍須熟悉句面層次的正確發音。

關鍵詞：音高重音、第二語言習得、日語、台灣華語