

## Against Marking Accent Locations in Japanese Textbooks

HASEGAWA Yoko\*

Key words: Japanese, accent, phonology, phonetics, intonation

Most textbooks of introductory Japanese mention that Japanese is a pitch-accent language, where every syllable in a word is either high- or low-pitched, and accent location is signaled by the last high-pitched syllable; e.g., *tocho'kan* 'library' has the L-H-L-L pitch configuration. They also mention that high vowels are devoiced in certain phonological environments; e.g., *kusa* 'grass' is pronounced as [k<sup>w</sup>sa]. These two characterizations result in pronunciations which are unattainable, namely high-pitched devoiced vowels; e.g., *shi'ki* 'four seasons' has the H-L pattern. The fact that native listeners do hear an accent on a devoiced syllable indicates that associating an accent invariably with a high pitch cannot be an accurate description of the language. This paper discusses how Japanese accent is actually realized and argues that marking accent locations in textbooks without a detailed explanation about accent is merely an extra complication that introductory textbooks should avoid.

### INTRODUCTION

Most Japanese textbooks briefly explain that the Tokyo dialect of Japanese (sometimes called "Standard Japanese") is a pitch-accent language in which accent location is signaled by the last high-pitched syllable, e.g., *tabe'ru* 'eat', and *nihongo no ji'sho* 'Japanese dictionary'.<sup>1</sup> Typically, the first syllable is low pitched if it is not accented; the following syllables are high pitched up to the accented syllable if there is one; and all subsequent syllables are low pitched. Textbooks commonly represent such pitch patterns as shown in Figure 1.

These textbooks also mention that the high vowels *i* and *u* are devoiced in certain phonological environments: typically between voiceless obstruents (stops and

\* 長谷川葉子: Assistant Professor, Department of East Asian Languages, University of California, Berkeley, U.S.A.

<sup>1</sup> The notation *tabe'ru* is used in this paper to indicate that /be/ is lexically accented. Different notations such as *tabe'ru*, *tabéru*, or *tabèru*, are used in some textbooks.

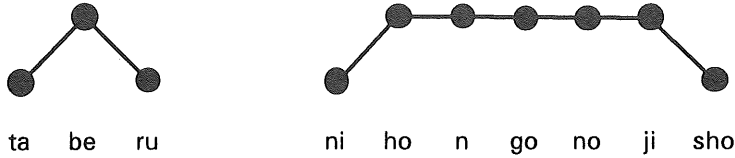
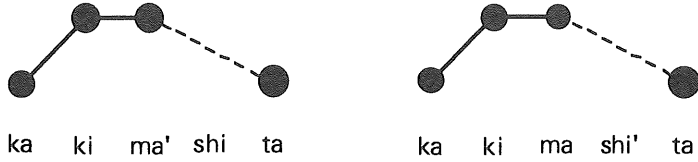


Fig. 1. Common model of Japanese pitch accent

Fig. 2. Pitch accent of *kakima'shita* and *kakimashi'ta*

fricatives) or in sentence-final position. For example, *shikei* 'capital punishment' and *tabemasu* 'eat' (polite) are normally pronounced as [ʃke:] and [tabemas].<sup>2</sup> In principle, these characterizations are accurate, but when applied simultaneously, they can result in confusing descriptions of the language, as observed in such textbooks as *An Introduction to Modern Japanese* and *Japanese: The Spoken Language*, which both mark accent locations extensively.

This paper argues that marking accent locations throughout a textbook, without further explanation of how pitch accent is realized in Japanese, does not aid, but can hinder, good pronunciation.

### Problems

Consider the following textbook description: "Accent in Japanese is a pitch accent in contrast to the stress accent in English. Each word in Japanese has a set accent, that is, certain syllables have a high pitch and others have a low pitch. We say that syllables within a word are either high or low, but this is a matter of *relative* rather than absolute pitch" (*Introduction to Modern Japanese*, pp. xiv–xv, emphasis in original).

Consider, then, words like *kakima'su* 'write' and *kakima'shita* 'wrote'. Because the final syllable *su* in *kakima'su* and the penultimate *shi* in *kakima'shita* are typically devoiced, their pitch cannot be lower than that of the preceding syllable. That is, they are acoustically and auditorily the same as *kakimasu'* and *kakimashi'ta* (see Figure 2).

More problematic examples are found when a lexically accented syllable is devoiced, e.g. *shi'ki* 'four seasons', *hashit'teiru* 'be running'. Because the *shi* in the former and the *shit* in the latter are voiceless, they cannot be higher than other

<sup>2</sup> The International Phonetic Alphabet is used to represent pronunciation.

syllables in pitch. Textbooks should avoid such self-contradictory descriptions.

On the other hand, most native speakers believe that an accent falls on *shi* in *shiki* or *shit* in *hashitteiru*, even though they are not uttered with a high pitch. In fact, if native speakers are forced not to devoice these vowels, they pronounce the words with a tonal configuration of high-low and low-high-low, respectively. Furthermore, even when *shi* in *shiki* is devoiced, they tend to believe that they have heard it high-pitched. How is this belief possible?

### Perception of Pitch Accent in Japanese

The claim that Japanese is a pitch-accent language is based almost exclusively on native speakers' introspection or impressionistic data. Onishi (1942) argued that because the function of accent is to differentiate the meaning of, or to make prominent a portion of, a word or phrase, any set of features that can serve these purposes (e.g., loudness and duration) may be distinctive. In the case of Japanese, he suggested that accent was an impressionistic sum of pitch and loudness.

Neustupný (1966) found positive evidence for Onishi's claim. Analyzing Japanese words by mechanical means, he measured the contours of fundamental frequency (abbreviated as F0)<sup>3</sup> and amplitude<sup>4</sup> for each word in his data. If Japanese were in fact a pure pitch-accent language, then F0 would fall around the boundary between an accented syllable and the following syllable. He found, however, that accent, as conventionally known, and the real F0 fall often do not synchronize: F0 fall may be delayed in relation to an accented syllable. He called this phenomenon *oso-sagari* (delayed F0 fall). He thus claimed that the F0 data by themselves are not sufficient for determining the accent pattern, and that since, in his data, the amplitude peak fell on the accented syllable in the words in which the F0 fall was delayed, both F0 and amplitude are distinctive features in the Japanese accentual system. In other words, he claimed that Japanese is both a pitch- and stress-accent language.

On the basis of acoustic and perceptual experiments using synthetic speech, Sugito (1972, 1982) refuted the Onishi-Neustupný hypothesis. She found that native Japanese listeners perceive an accent on a syllable when the syllable is followed by a falling F0 contour, even though the F0 peak of the accented syllable is no higher than the following syllable.

To support Sugito's discovery, Hasegawa and Hata (1988) presented the following pair of tokens from their production data.

<sup>3</sup> Fundamental frequency (F0) is an objective property of sound, which can be measured by a mechanical device. Pitch, on the other hand, is a sensation, and thus, is subjective by nature. Generally, we hear a sound as high-pitched when its F0 is high. However, machines can be more sensitive than the human ear, while our ears are not uniformly sensitive to all sounds. We cannot, therefore, consider F0 (physical property) and pitch (subjective judgment of the sound) one and the same thing. Some may hear two sounds at the same pitch, but others may hear them differently. See Ladefoged (1962) for details.

<sup>4</sup> Similar to the F0/pitch distinction, amplitude is a physical property of sound, whereas loudness is a subjective judgment of sound.

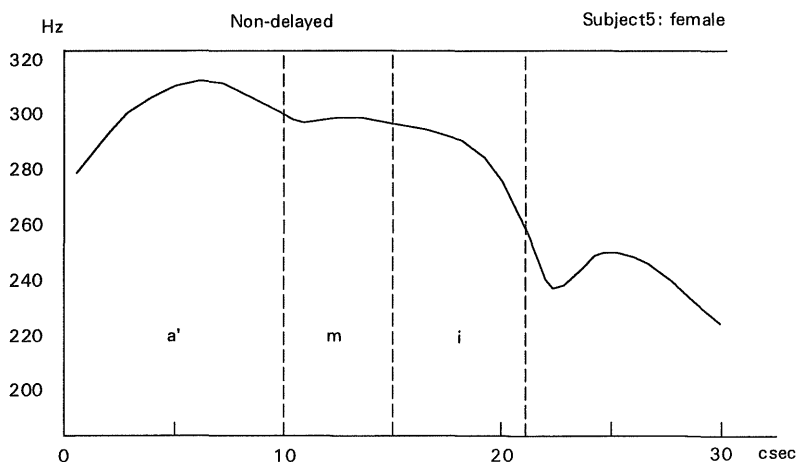


Fig. 3.1. F0 contour for the word *na'mida* (non-delayed token)

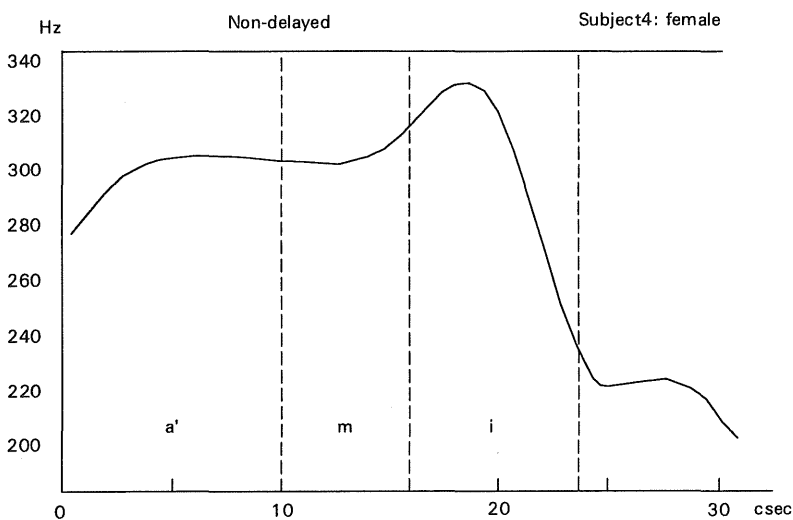


Fig. 3.2. F0 contour for the word *na'mida* (delayed token)

These figures are part of the word *namida* 'tear' (Noun) (uttered by two female native speakers of Japanese), in which the lexical accent falls on the first syllable. In Figure 3.1, the F0 peak is on *a*, which is in accordance with the lexical accent location. In Figure 3.2, by contrast, the F0 peak is clearly on *i*, and yet the word was perceived as *na'mida*. As Neustupný reported, F0 fall sometimes delays with respect to the accented syllable, without listeners detecting such a delay. Figure 4 schematically shows this relationship between delayed F0 fall and perceived accent.

This phenomenon of illusory pitch accent explains why native listeners perceive an accent on a devoiced vowel. Even though a high F0 cannot occur on a perceived vowel, the F0 fall on the following syllable forces native listeners to associate an

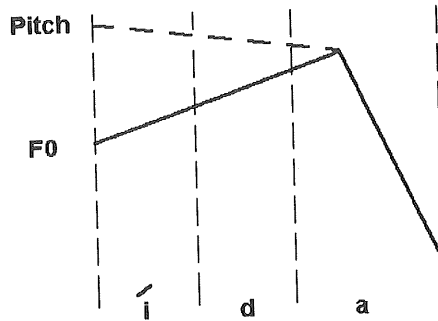


Fig. 4. Perceived accent and the actual F0 peak

accent with the preceding syllable containing the devoiced vowel.

The question of whether non-native speakers of Japanese perceive such stimuli as Figure 3.2 in the same way as native speakers do has direct relevance to the teaching of Japanese. It has been reported that some native speakers of English do, and others do not (Hata and Hasegawa, 1991; Hasegawa and Hata, 1992). This implies that perceiving the stimulus in Figure 3.2 with a high pitch on the first syllable is a learned response, and thus is not based on a universal constraint of human auditory perception. I would predict that when the first syllable contains a devoiced vowel, e.g., *ki'kai* [k<sup>h</sup>kai] 'chance' (unlike *namida* in which the first vowel is voiced), most, if not all, non-native speakers do not perceive the first syllable as accented — even if it is followed by an F0 fall.

### Discussion

Sometimes what people believe they heard and what they actually heard are different. For example, most native speakers of English believe that the difference between *Chi* (a Greek letter) and *guy* is such that the first consonant is voiceless in the former and voiced in the latter. Phonologists contend, however, that this is not the case. The first consonants in both words are voiceless, and they are differentiated only by aspiration: the first consonant is aspirated in the former, while it is unaspirated in the latter. On the other hand, when the consonant /k/ appears after an /s/, it is normally not aspirated in English. Therefore, if we record the words *Chi*, *guy*, and *sky* spoken by a native speaker of English and splice the tape, the *sky* without the /s/ sounds like *guy*. Also, if we add the /s/ portion from the *sky* in front of the *guy*, the new word sounds like *sky*.<sup>5</sup>

Another example of the discrepancy between native speakers' belief and their actual auditory perception involves the so-called stress accent of English. Most speakers believe that accent in English is manifested by loudness due to the term *stress*. Fry (1958) conducted perceptual experiments with synthetic noun-verb pairs

<sup>5</sup> This experiment is due to John Ohala.

in which the distinction is made only by the accent placement, e.g., *súbject* (noun) vs. *subjéct* (verb). He found that the increase in vowel duration of the second syllable can cause a perceived accent shift from noun *súbject* to verb *subjéct*. The increase in amplitude has a similar effect, although to a lesser degree. As for the ranking between F0 and duration cues, typically the former outweighs the latter. Therefore, in Fry's experiment, the most significant cue of English accent is pitch, followed by length, and then loudness.

Native speakers of Japanese display similar discrepancies in auditory perception. One salient example is that Japanese pitch accent is most accurately represented in such models as in Figure 1. As pointed out in the previous section, this model cannot account for the perception of accent on a syllable with a devoiced vowel. In fact, Japanese accent is similar to English accent, as Beckman and Pierrehumbert (1986) and Pierrehumbert and Beckman (1988) have convincingly argued. Only a few syllables in Japanese are associated with a high or low pitch. Beckman and Pierrehumbert claim that a syllable bears a high or low pitch only when it is in word or phrase initial/final position, in a second position in a word or phrase, or when it is lexically accented.<sup>6</sup> The difference between English and Japanese accent is that Japanese only has pitch as a psychoacoustic cue (Fujisaki and Sugito, 1977), while English has a number of cues (i.e., length, loudness, and vowel quality) that influence the perception of accent (Beckman, 1986).

It is frequently observed that in order to make students mark accent locations correctly, the instructor of Japanese uses cues such as loudness, which are not conventional in Japanese. And it is almost always the case that instructors read words with unnatural, stepwise high and low pitch, e.g., *to(L)sho(H)ka(L)n(L)*, when they test students' ability to detect pitch accent. Through this kind of training, students cannot improve their ability to detect accentual patterns in naturally uttered sentences, because no native speaker speaks Japanese in such a way. This is why I believe that marking accent in textbooks is at best an extra complication, and at worst can cause adverse effects if instructors and students take it seriously.

### How to Teach Japanese Accent Then?

Recently, I read in *Mangajin* (No. 20, September 1992, p. 20) a letter from a learner of Japanese about pitch accent and the editor's reply, which is lucid and theoretically sound. I present both of them here.

**Letter to the Editor:** I've grown acutely accent-sensitive by studying [Eleanor] Jordan's textbook and I miss the marking of accents on MANGAJIN's romanization. If the idea of the magazine is to make the material pretty much self-contained for language learning, it doesn't work quite right in my case, because without resorting to a dictionary it would feel as if (well, not quite, but . . .) I were learning the spelling of English words without caring about

---

<sup>6</sup> I have simplified Beckman and Pierrehumbert's claims for expository purposes.

the pronunciation. In the “pronunciation guide” you dismiss the intonation as mostly inessential, and the majority of the *kokugo*, *ei-wa*, or *wa-ei* dictionaries back up your view by simply ignoring accent. . . . I would just feel vindicated by an acknowledgment that accent, *yappari*, is an issue.

**Reply from the Editor:** We would not deny that “accent is an issue,” but we think that imitating native speakers, whether in real life or on the tapes that go with your textbook, is more likely to produce natural-sounding results than attempting to fabricate the sound on your own from a notation or explanation given in writing. This is true for all matters involving pronunciation, which is exactly why our pronunciation guide has the disclaimer you mention. . . .

It’s interesting to note that native Japanese speakers outside Tokyo speak otherwise standard Japanese (*hyoojungo*) with different “pitch accents” (this is what we are speaking of here, not dialect accents) and never have trouble being understood. For the student of Japanese, a flat, even intonation will always be understood, and for Americans (and some Europeans) who tend to give their words very marked pitch accents, this may be a good way to eliminate some un-Japanese sounding speech habits.

When two or three words sound exactly alike except for pitch accent, context is going to resolve the ambiguity virtually 100 percent of the time. In practical terms, accent is probably the least important aspect of Japanese pronunciation no matter what your level of language skill.

On the whole, we think most people are best off following Jack Seward’s advice: “the degree of variance in pitch is so small that the beginner is advised to voice all Japanese words . . . with a steady evenness of pitch . . . Sooner or later, depending on the sharpness of your ear, you will come to be able to distinguish among and mimic the existing minor variations in pitch.”

I wholeheartedly support this advice. I do not, however, claim that teaching of Japanese accentual patterns should be avoided *in toto*. On the contrary, I believe that students should be informed about Japanese accent during the early stages of learning. It is likely that the best way is to make students realize the major differences in F0 modulation between Japanese and their native languages. For native speakers of English, whose language has an accentual system similar to that of Japanese, it suffices to mention that they avoid making accented syllables louder and/or longer. For native speakers of a tone language, e.g., Chinese, Thai, and Vietnamese, it should be mentioned that not all syllables, but only a few in Japanese are associated with tones.

## CONCLUSION

Many Japanese textbooks presuppose that all syllables in the language are either high-pitched or low-pitched, and some mark accent locations extensively. However, recent studies show that not all syllables in Japanese are associated with a particular pitch. Furthermore, if the commonly used notations in accent marking are taken

seriously, those textbooks run into problems because an accent may fall on a syllable with a devoiced vowel — which is physically impossible.

One might argue that the discrepancy I discussed in this paper does not alone provide sufficient justification for eliminating accent markings from introductory textbooks because phonological representations and their surface phonetic realizations need not be identical. For example, many, if not most, English vowels are centralized in natural speech, but we still need to record the “underlying” or “idealized” vowel quality somewhere. I would argue that abstract phonological notations may be recorded in dictionaries, but not in textbooks. Furthermore, the students must be told that pronunciation markings in dictionaries are not what they are supposed to hear or utter in natural running speech.

People without hearing impairments can mimic the melody of language, but they can hardly interpret visual accent markers into the oral/aural domain without special training because visual and auditory stimuli are processed very differently in the human brain. In all likelihood, the author of the above-mentioned letter simply feels more comfortable *visually* with accent markers. But using such markers to speak Japanese creates pronunciations that are worse than a crude synthesizer.

#### BIBLIOGRAPHY

- Beckman, Mary E. 1986. *Stress and non-stress accent*. Dordrecht, Holland: Foris.
- Beckman, Mary E. and Janet B. Pierrehumbert. 1986. Intonational structure in English and Japanese. *Phonology Yearbook* 3: 255–309.
- Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and Speech* 1: 126–52.
- Hasegawa, Yoko and Kazue Hata. 1988. Delayed pitch fall in Japanese. *Journal of the Acoustical Society of America*, Suppl. 1.83: S29.
- . 1992. Fundamental frequency as an acoustic cue to accent perception. *Language and Speech* 35: 87–98.
- . 1988. Delayed pitch fall in Japanese: A perceptual experiment. *Journal of the Acoustical Society of America*, Suppl. 1.84: S156.
- . 1991. The effect of F0 fall rate on accent perception in English. *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, 121–29.
- Jorden, Eleanor H. and Mari Noda. 1987–90. *Japanese: The spoken language*. New Haven: Yale University Press.
- Ladefoged, Peter. 1962. *Elements of acoustic phonetics*. Chicago: University of Chicago Press.
- Mizutani, Osamu and Nobuko Mizutani. 1977. *An introduction to modern Japanese*. Tokyo: Japan Times.
- Neustupný, J.V. 1966. Is the Japanese accent a pitch accent? *Onsei-gakkai Kaihō* 121. Reprinted in *Akusento*, ed. M. Tokugawa. Tokyo: Yūseidō, 1980.
- Pierrehumbert, Janet B. and Mary E. Beckman. 1988. *Japanese tone structure*. Cambridge, Mass.: MIT Press.



杉籐美代子 (1972) 「おそ下がり考——動態測定による日本語アクセントの研究 (その一)」、大阪樟蔭女子大学論集 10.

———. (1982) 『日本語アクセントの研究』、三省堂.

藤崎博也、杉籐美代子 (1977) 「音声の物理的性質」、『岩波講座日本語 5: 音韻』、岩波書店.