

## Prosodic Features of Japanese and English

Marie KJELDGAARD

### 要 旨

日本語と英語のプロソディーの概要を紹介する。言語学におけるプロソディーとは、発話において現れるリズム、音量、スピード、イントネーション等をいう。「ToBI」という音訳システムを用いて、英語と日本語のプロソディーの共通点と相違点を示す。

Keywords: English (英語), intonation (イントネーション), Japanese (日本語), language differences (言語の違い), linguistics (言語学), prosody (プロソディー)

### Abstract

This paper contains an overview of the prosody of Japanese versus that of English. Prosody refers to the rhythm, volume, tempo, and intonation of connected speech. This article also gives an overview of the ToBI transcription system for recording prosodic features. Aspects of both English and Japanese prosody are described, with the goal of pointing out both similarities and differences between the two languages.

## Introduction

What are prosodic features? According to Celce-Murcia, Brinton, and Goodwin (1996), prosodic features include rhythm, volume, tempo, and intonation in connected speech. Rhythm comes from the use of stressed (or accented) and unstressed sounds, such as the pattern of stressed and unstressed syllables found in many Germanic and Romance languages. Volume refers to prominence (loud or soft), and tempo could be described as the length of a sound (long or short). Intonation corresponds to pitch (high or low).

These linguistic features can correspond roughly to physical, measurable quantities such as intensity in decibels, duration, and fundamental frequency ( $F_0$ ), or pitch in Hertz; however, these features are not always present or do not always match exactly, especially in cases of tonal languages such as Chinese (Hirst & Di Cristo, 1998). There is also some ambiguity in the literature in the use of the terms *intonation* and *prosody*. Intonation is often used in a broad sense, meant to include the other aspects of prosody; other times it is used more narrowly, to refer only to intonation proper (Hirst & Di Cristo).

Prosody is often not taught explicitly or effectively in the language classroom (Allen, 1971; Levis, 1999), but it is an important part of language production. In fact, according to Celce-Murcia et al. (1996), misunderstandings resulting from improper intonation use can be more difficult to repair than those resulting from segmental pronunciation errors. Even if a speaker uses clear and correct grammar, a mistake in intonation can change the meaning or affective content of an utterance. Prosodic features are such a fundamental part of language that studies have shown babies can recognize the intonation of their language as early as two months of age, before they acquire lexical items (e.g., Fernald 1989; Jusczyk, Cutler, & Redanz, 1993).

Considering these points, it seems important for language teachers to have an understanding of the prosodic features of the language they are teaching; knowledge of the student's L1 prosody would also be helpful, allowing the teacher to identify potential areas of difficulty. This paper provides a brief overview of some of the similarities and differences of prosodic features in English and Japanese.

## **A Note on Transcription**

Because prosodic features consist almost entirely of auditory information, there have been some difficulties in deciding how they should be represented on paper. Hirst and Di Cristo (1998) explained several different methods of intonation transcription, used in different chapters of their book. Beckman and Pierrehumbert (1986), in their influential description of Japanese and English intonation, used a relatively simple system which marked high and low tones, accents, and breaks in speech. They later worked with several other authors to develop that system into a more widely useable form, resulting in the Tones and Break Indices (ToBI) intonation notation system (Silverman et al., 1992). This system is particularly useful because it is adaptable to other languages: A version specifically used to describe Japanese intonation, called J\_ToBI, has been developed (Venditti, 2005). Therefore, the ToBI and J\_ToBI notation styles are referenced in this paper.

## **Intonation in Japanese and English**

### **Accents and Tones**

In prosody, accents are abrupt pitch changes which occur on the stressed syllable of a word. (Bolinger, 1998, p.47). Tones refer to changes in pitch that carry lexical meaning. Languages are often classified according to their use of accents, tones, or both. In tonal languages—for example, Chinese, Vietnamese, or Thai—pitch is used to distinguish between multiple lexical items. This means that two words with the same pronunciation will convey different meanings depending on the pitch or tone. Accent is usually not found in tonal languages.

Stress languages (sometimes called stress-accent or intonation languages) do not use pitch to distinguish between two different words. Instead, pitch is used to convey intonational information. Most Germanic and Romance languages, including English, fall into this category.

There are also languages, like Japanese, that fall somewhere in the middle; they are called pitch-accent or tonal accent languages. Tone can convey lexical meaning; however, it is not always used in this way. In some words, tone is anchored to one accented or stressed syllable in a word, and the other syllable or syllables are unaccented. In all other

words, the entire word is unaccented (Hayes, 2009).

Hirst and Di Cristo (1998), however, warn against considering these three categories as absolute. They suggest that, rather than being distinct, they could be seen as points on a scale, with individual languages falling somewhere along that scale. Table 1 shows ways that accent and pitch are used in three different languages.

**Table 1 Stress and Tone in Chinese, Japanese, and English**

Language	Type	Uses accent?	Uses tone for:	Example
Chinese	tone	no	lexical meaning	<i>mā</i> (high tone) means “mother” <i>mǎ</i> (falling-rising tone) means “horse”
Japanese	pitch-accent	some words	lexical meaning in some words	<i>káki</i> (accent on first syllable, drop in pitch on second syllable) means “oyster” <i>kaki</i> (no accent, rise in pitch on the second syllable) means “persimmon”
English	stress	yes	intonational meaning	<i>Red.</i> (falling intonation) is a statement of fact <i>Red?</i> (rising intonation) is a question or expression of surprise

*Note.* Partially adapted from Abe (1998).

### Pitch-Accents

Pitch accents occur in English as well as in Japanese. A pitch accent refers to a high or low tone that is associated with a stressed syllable. In their initial research on Japanese and English intonation, Beckman and Pierrehumbert (1986) identified six different kinds of pitch accents in English. Upon development of the ToBI transcription method, the original six were reduced to five (Pitrelli, Beckman, & Hirschberg, 1994). They identified just one in Japanese. Table 2 shows a list of possible pitch accents in English and in Japanese.

In English, pitch accents provide intonational meaning. Different pitch-accents can be used with the same words. In Japanese, the pitch accent has no inherent intonational meaning. It is associated only with specific words (Venditti, 2005). Pitch accents in Japanese are realized with a single intonation pattern; all pitch-accents in Japanese follow this same H to L tone pattern. It is important to note, however, that Japanese does have

many of the same intonational contours as English. In English, they are associated with a pitch accent; in Japanese, they are not.

Another important difference is the frequency of occurrence. In English, every phrase is built around a pitch accent, which marks the key point of the phrase. In Japanese, pitch accents only occur in specific words. It is possible to have a phrase without any accents at all (Beckman & Pierrehumbert, 1986).

**Table 2 Pitch Accents in English and Japanese**

English		
Type	Intonation Pattern	Example of Usage
H*	simple high	canonical declarative
L*	simple low	yes-no question
L + H*	rising to high from low	contrastive focus
L* + H	“scooped”	pragmatic uncertainty
H + !H*	fall onto stress	pragmatic inference
Japanese		
Type	Intonation Pattern	Example of Usage
H* + L	sharp fall	indicates lexical meaning

*Note.* Adapted from Pitrelli, 1994 and Venditti, 2005. Asterisks indicate association with a stressed or accented syllable.

### Rhythm and Stress

Japanese and English also differ in their rhythm, or timing. English is a stress-timed language, in which the length of an utterance depends more on the number of stresses, rather than the number of syllables. Figure 1 gives an example of this.

The CATS	CHASE	MICE.	(4 syllables)
The CATS will	CHASE the	MICE.	(6 syllables)
The CATS have been	CHAS-ing the	MICE.	(8 syllables)

**Figure 1. Example of stress-timed language. Adapted from Celce-Murcia et al., 1996.**

Although each sentence contains a different number of syllables, all three sentences contain the same number of stresses (on *cats*, *chase*, and *mice*, indicated by capital letters). Therefore, they should all take roughly the same amount of time to say, even though the third sentence has twice as many syllables as the first.

Japanese, on the other hand, is a syllable-timed—or, technically, a mora-timed—language. A mora is a subsyllabic unit; it is “a postvocalic segment within a syllable, such as a coda consonant or the second element of a diphthong, [which] functions independently as a rhythmic unit” (Kureta, Fushimi, & Tatsumi, 2006, p. 1102). Syllables in Japanese contain either one or two morae; for example, the word *Nippon* (“Japan”) contains four morae (*ni-p-po-n*) but just two syllables (*nip-pon*).

In English, strong syllables—usually those containing new or important information—are stressed; weak syllables do not receive stress. In Japanese each mora receives roughly equal weight. Therefore, the length of a phrase in Japanese will increase in proportion to the number of morae in the phrase. Figure 2 shows an example of this. The first phrase contains 9 morae; the second has 18, so it will take about twice as long to say.

<p>猫 は ネズミ を 追う。  <i>Ne/ko / wa / ne/zu/mi / wo / o/u.</i> (9 syllables)  Cats chase mice.</p> <p>その 猫 は その ネズミ を 追いかけている。  <i>So/no / ne/ko / wa / so/no / ne/zu/mi / wo / o/i/ka/ke/te/i/ru.</i> (18 syllables)  That cat is chasing that mouse.</p>
--

**Figure 2. Example of mora-timed language. Adapted from Ohata, 2004.**

Stress-timed languages like English can pose difficulties for learners with a syllable- or mora-timed L1. They may tend to stress syllables equally, without reducing weak syllables or stressing strong syllables sufficiently. According to Ohata (2004), stress in Japanese comes from simply saying the word at a higher pitch, whereas in English, stressed syllables are also longer and louder. This difference contributes to the rhythmic pattern of stresses in English, as well as the more uniform sound of Japanese: changing just the pitch does not lend the same prominence to stressed syllables as changing pitch, intensity, and duration. In addition, many Japanese words have no accent, so it is not uncommon for entire phrases to be unaccented (Beckman & Pierrehumbert, 1986).

### **Pitch and Intonational Patterns**

The pitch range used in Japanese versus in English also tends to vary; in general, Japanese uses less pitch variation, while English uses a wider range (Han, 1963; Ohata,

2004). In fact, Han claimed that “what is a significant rise or fall in Japanese does not sound important at all to the English speaker’s ear” (p. 8). He stated that in Japanese, the relative change in pitch between syllables is more important than the absolute range. Although Japanese and English both use some similar intonation patterns—for example, the use of rising intonation to indicate a yes-no question—the degree of change in pitch tends to be different (Ohata, 2004).

Pitch is measured over the course of a contained unit of speech often called an intonation unit or intonational phrase (IP). An IP consists of smaller units known as accentual phrases (APs). An IP could be as short as a single word, or it could cover an entire multi-word sentence. In English, it should be marked by pauses both before and after; it should be grammatically cohesive; it should contain (at least) one prominently stressed element; and it should contain a complete intonation pattern (Celce-Murcia et al., 1996).

Intonational meaning is created at the AP level. In Japanese, the AP is also the realm of a process known as *downstep* or *catathesis*, where a pitch-accent triggers a falling tone. Catathesis can be repeated, creating a “descending staircase” effect over multiple APs (Beckman & Pierrehumbert, 1986; Pierrehumbert & Beckman, 1988). Catathesis can occur in both Japanese and English.

The IP is the domain of *final lowering*, where the pitch is lowered significantly at the end of a declarative utterance; this also can occur in both languages. *Declination*, in which the fundamental frequency of the pitch gradually lowers—or declines—over the course of the utterance, also occurs on the IP level in Japanese (Beckman & Pierrehumbert, 1986; Deguchi & Kitagawa, 2002; Pierrehumbert & Beckman, 1988). Pierrehumbert & Beckman did not find evidence of declination in English.

## Summary

Although there are some similarities, prosody in English and Japanese does exhibit many differences. These could potentially pose problems for Japanese learners of English. Items that showed the most variation included the use of accents to provide intonational meaning in English, while they are used less frequently as lexical item markers in Japanese; the difference between stressed and unstressed syllables in English and timed mora in Japanese; and the greater range of pitch in English versus a reduced pitch range

and declination in Japanese. These, then, are the prosodic features that are important to be aware of as a language teacher. They should also be introduced to learners, in order to aid them in comprehension and production of the language.

## References

- Abe, I. (1998). *Intonation in Japanese*. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 360–375). Cambridge, MA: Cambridge University Press.
- Allen, V. F. (1971). Teaching intonation, from theory to practice. *TESOL Quarterly*, 5(1), 73–81.
- Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255–309.
- Bolinger, D. (1998). *Intonation in Japanese*. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 360–375). Cambridge, MA: Cambridge University Press.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge, MA: Cambridge University Press.
- Deguchi, M., & Kitagawa, Y. (2002). Prosody and wh-questions. In M. Hirotsu (Ed.), *Proceedings of the thirty-second annual meeting of the North East Linguistic Society* (pp. 73–92). Amherst: GLSA.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 60(6), 1497–1510.
- Han, M. (1963). The problem of pitch in the teaching of Japanese. *The Journal-Newsletter of the Association of Teachers of Japanese*, 1(1), 5–9.
- Hayes, B. (2009). *Introductory phonology*. West Sussex, UK: Wiley & Blackwell.
- Hirst, D., & Di Cristo, A. (1998). *A survey of intonation systems*. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 360–375). Cambridge, MA: Cambridge University Press.
- Jusczyk, P., Cutler, A., & Redanz, N. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, 64(3), 675–687.
- Kureta, Y., Fushimi, T., & Tatsumi, I. (2006). The functional unit in phonological encoding: Evidence for moraic representation in native Japanese speakers. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 32(5), 1102–1119.
- Levis, J. (1999). Intonation in theory and practice, revisited. *TESOL Quarterly*, 33(1), 37–63.
- Ohata, K. (2004). Phonological differences between Japanese and English: Several

- potentially problematic. *Language Learning*, 22, 29–41.
- Pierrehumbert, J., & Beckman, M. (1988). *Japanese tone structure*. Cambridge, MA: The MIT Press.
- Pitrelli, J., Beckman, M., & Hirschberg, J. (1994). Evaluation of prosodic transcription labeling reliability in the ToBI framework. *Proceedings of the International Conference on Spoken Language Processing, Yokohama, 2*, 123–126.
- Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). TOBI: a standard for labeling English prosody. In *The Second International Conference on Spoken Language Processing, ICSLP 1992, Banff, Alberta, Canada, October 13–16, 1992*.
- Venditti, J. J. (2005). The J\_ToBI model of Japanese intonation. In S. A. Jun (Ed.) *Prosodic typology: The phonology of intonation and phrasing* (pp. 172–200). Oxford: Oxford University Press.