

THE PHONOLOGY AND PHONETICS OF SECOND LANGUAGE INTONATION: THE CASE OF “JAPANESE ENGLISH”

Motoko Ueyama

Department of Applied Linguistics & Phonetics Laboratory at the Department of Linguistics, UCLA
405 Hilgard Avenue Los Angeles, CA 90025-1543, U.S.A.
(ueyama@humnet.ucla.edu)

ABSTRACT

A production experiment was conducted in order to examine the acquisition of English intonation by native speakers of Japanese, and the results were analyzed within the framework developed by Pierrehumbert [3] and her colleagues. The results suggest that second language intonation is acquired on two different levels: learners first acquire the categorical patterns of the foreign intonation, and only later learn to produce native-like continuous intonational streams. This supports models in which the speech cognitive system is split into two sub-modules: a phonological component (characterized by categorical units) and a phonetic component (implementing the phonological units as a continuous articulatory/acoustic stream).

1. INTRODUCTION

1.1. Two-Module Intonation Models

Many studies explicitly or implicitly assume a model in which the speech system is split into two modules: the phonological and the phonetic component. The phonological component is characterized by a set of language-specific categorical units. The phonetic component “implements” these symbolic units to generate/process the continuous articulatory/acoustic stream [1, 2]. A classic example of this approach to speech is the two-level model of intonation first formulated explicitly by Pierrehumbert [3]. Within her framework, continuous intonation contours are analyzed as sequences of two discrete phonological units (H and L tones), that are phonetically interpolated.

At the *phonological level*, sequences of abstract tones are specified, reflecting the hierarchical structure of phrasing. In English, two hierarchical phrasing units that are higher than word level are postulated [3, 4]: the

Intonational Phrase (IP) and the intermediate phrase. The IP is the largest phrasing unit, i.e. the highest layer of the prosodic structure. The IP consists of one or more intermediate phrases. Each intermediate phrase contains one or more pitch accents. In this model, underlying phonological tones are categorized as one of three types: pitch accents, phrasal tones, and boundary tones. The pitch accent is a tonal target associated with a stressed syllable (H*, L*, H+L*, H*+L, L+H*, L*+H). The boundary tone marks the end of the IP (L%, H%). The phrasal tone fills in the space between the last pitch accent and the boundary tone (L-, H-).

At the *phonetic level*, the underlyingly specified sequences of abstract tones (i.e. fundamental frequency/f₀ targets) are interpolated, following certain regularities. Crosslinguistically common regularities are declination (i.e. overall f₀ downtrends towards the end of the utterance) and downstep (i.e. a stepwise lowering of f₀ at points underlyingly associated with high tones in the utterance), as reviewed by Ladd [5].

1.2. Focus Realization in Interrogatives in English and Japanese

The production of English contrastive focus in interrogatives by Tokyo Japanese speakers was examined by Ueyama & Jun [6], adopting Pierrehumbert’s model of English intonation [3]. The patterns of focus realization in interrogatives of American English and Tokyo Japanese (henceforth, English and Japanese, respectively) are schematized in Figure 1. In both languages, the high boundary tone (H%) marks the end of the phrase. However, the two languages realize other properties of narrow focus differently. While (in interrogatives) a focused word in English is realized with a low tone (L*) followed by a high plateau (H-), a focused word in Japanese is realized with a high tone followed by a low plateau phonologically specified as HL [3, 4, 7]. Thus, the tonal contours of focus

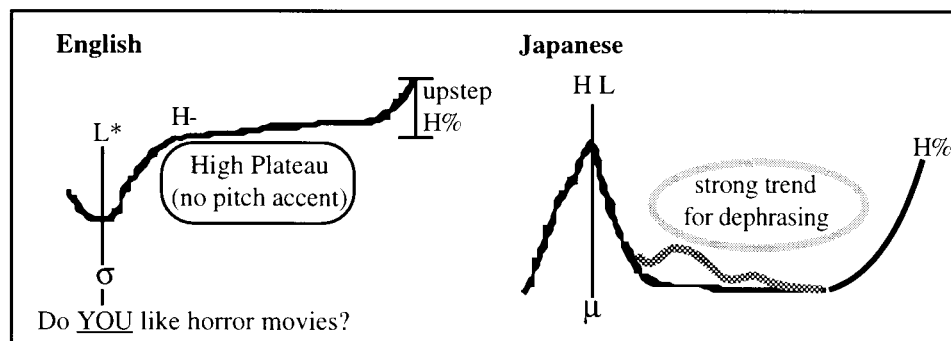


Figure 1. Schematic contours of focus realization in interrogatives

realization in interrogatives can be phonologically described as L* H- H% in English and HL H% in Japanese. The English high plateau means that pitch accents after focus are deaccented (phonological property) and are realized by sustaining high f0 (phonetic property). In Japanese, deaccentuation after focus is not obligatory but a strong tendency: the low plateau is often perturbed by sequences of downstepped pitch accents [4, 7, 8]. The two languages also differ in another phonetic property: the slope of initial f0 rise in the interrogatives. In English, after the focused word L*, f0 reaches its peak near the *end of the following word*. On the other hand, in Japanese, f0 after the phrase-initial L tone in the focused phrase reaches its peak near the *second syllable/mora of the phrase*, thus creating a sharper rising slope than that in English.

1.3. Goal

Based on the data on L2 English produced by Korean and Japanese speakers presented in [6], we claimed that L2 learners acquire the intonational system of the target language in two steps. First, they internalize the phonological structures of L2 intonation, while they still implement these structures by using L1 phonetic habits. Only later, learners master native-like phonetic regularities. In this study, I will investigate the validity of this claim on the basis of the data from [6] and additional data.

2. EXPERIMENT

2.1. Subjects

Two major speaker groups were compared: two native speakers of American English (1 male & 1 female) as a control group, and five Japanese learners of L2 English (2 males & 3 females) as an experimental group. Three proficiency levels were compared within the experimental group.

2.2. Corpus

Test sentences were designed to test different degrees of plateau length by changing the numbers of nouns in the object noun phrase (NP):

1. Do **you** need my Grandma's marmalade?
2. Do **you** need my Grandma's orange marmalade?
3. Do **you** need my Grandma's Mandarin-orange marmalade?
4. Do **you** need my Grandma's homemade Mandarin-orange marmalade?

The word in bold is focused. To trigger narrow focus, the interrogative sentences were embedded in the following frame:

None of my friends needs NP (3~6 words).

Do you need NP (3~6 words)?

2.3. Procedure

The order of sentences was randomized, and filler sentences were inserted pseudo-randomly in order to separate each target sentence. Subjects were recorded

in the sound booth at the UCLA Phonetics Laboratory. Each subject read the entire list 6 times. The speech data were digitized and the pitch tracks of the utterances were analyzed using Entropic's Xwaves+.

2.4. Measurements

For the phonological description of L2 intonation, the type of pitch accents and phrase boundaries occurring in each utterance was labeled adopting the framework described in [4]. Fundamental frequency (f0) and absolute time points were measured for the analysis of phonetic properties. As shown in Figure 2, each utterance was divided into two parts, *pre-plateau* and *plateau*, with reference to the center of the vowel in the verb "need". This verb was taken as the reference point since it is assumed to coincide with the beginning of the plateau following L* on the focused word [4]. Within the pre-plateau region, the lowest f0 and the highest f0 were measured together with their corresponding time values. Within the plateau region, the f0 and time value of the highest f0 point of each noun as well as the f0 time value of the utterance-final point were measured.

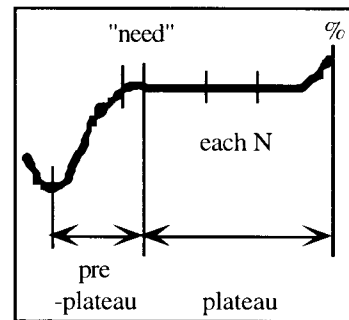


Figure 2. Schematic of measurement points

3. RESULTS & DISCUSSION

2.2. Phonology of L2 English Intonation

In order to investigate how the phonological aspects of focus realization in English are learned by Japanese speakers, the numbers of pitch accents and intermediate phrase boundaries were compared between the native and non-native speakers of English. The data of all female speakers (EF, JAF, JIF and JBF) are adopted from [6]. The results are summarized in Table 1. In this table, the star '*' indicates the number of pitch accents in each phrase and 'phr' refers to the number of phrase boundaries between the focused word and the boundary tone. Since there is no pitch accent after focus in English [3], we expect only one star tone and no phrase boundary in any sentence. Since each sentence is repeated 6 times, native English speakers should produce 6 stars and 0 phrase boundaries. Table cells showing this native-like distribution of tones and boundaries are shaded.

This distribution of pitch accents and phrase boundaries was found in the production by both native speakers of American English, EM and EF, along with

NP	EM		EF		JAF		JIM		JIF		JBM		JBF	
	*	phr	*	phr	*	phr	*	phr	*	phr	*	phr	*	phr
3 N	6	0	6	0	6	0	6	0	6	0	17	5	16	6
4 N	6	0	6	0	6	0	6	0	10	2	24	8	19	6
5 N	6	0	6	0	6	0	6	0	17	4	25	7	18	6
6 N	6	0	6	0	6	1?	6	0	22	5	19	13	19	6

Table 1. Results of tone type labeling. E = English; J = Japanese; M = male; F = female; A = advanced; I = intermediate; B = beginning

the realization of the expected tonal sequence L* H-H%. After the alignment of L* on the focused word “you”, no pitch accent was produced and the high plateau was realized as a sequence of H- and H%.

The advanced Japanese learner, JAF, and one intermediate learner, JIM, showed native-like patterns. This means that these two learners have mastered the phonology of focus realization in English interrogatives. The beginning learners, JBM and JBF, show a much higher number of pitch accents (*) and phrase boundaries, and this indicates that the two learners have not acquired phonological dephrasing yet. An intermediate stage is observed in the patterns of JIF. She successfully realized dephrasing after narrow focus in the shortest sentence, but produced more pitch accents and phrase boundaries as the length of the NP increased.

2.2. Phonetics of L2 English Intonation

Two learners, JAF and JIF, acquired the phonology of English focus realization in interrogatives. Does this mean that the acquisition of focus realization is completed? When the data are further examined with respect to phonetic properties, a much sharper contrast between the native speakers and even the more advanced Japanese learners emerged. This is particularly clear with respect to the realization of two phonetic properties of focus realization in interrogatives: (1) the sustenance of F0 in the high pitch range during the plateau part; (2) the initial F0 rise from the lowest pitch to the initial rise in the pre-plateau region.

2.2.1. F0 Sustenance in the High Pitch Range

Pitch tracks of representative speakers are shown in Figure 4. Both native speakers interpolated H- and H% by sustaining f0 in the high pitch range, as shown by EM’s pitch track. On the other hand, this pattern was not found in the speech of JAF and JIM, the learners with native-like phonology. As shown by JIM’s pitch track, F0 after L* is not sustained even though there is no extra pitch accent after focus. The three Japanese learners with non native-like phonology, JIF, JBM and JBF, lowered F0 immediately after focus and maintained f0 in the low pitch range, as shown by JBM’s pitch track. This difficulty in sustaining f0 in the high pitch range is probably due to the physical difficulty in holding high pitch for a long time period, and could be a universal constraint on L2 intonation acquisition in general.

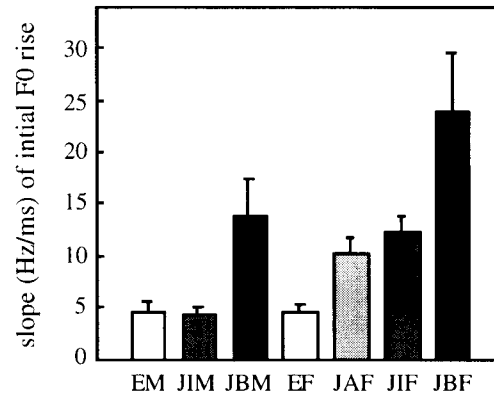


Figure 3. Slope of initial F0 rise

2.2.2. Initial F0 Rise

The slope of initial f0 rise is much steeper in Japanese than in English, as discussed earlier, and it is reasonable to expect that this phonetic property of Japanese will negatively transfer to English focus realization in interrogatives produced by Japanese learners. This prediction is supported by the data in Figure 3 on the slope of the initial F0 rise for each speaker (slope data for all female speakers are adopted from [6]). Here, short bars symbolize shallow slopes and tall bars symbolize sharp slopes. Generally, the slope is shallower for the more native-like speakers. These results suggest that the sharp slope in L1 pitch rise is negatively transferred, and that the degree of negative transfer decreases as the proficiency level increases. Notice that the absence of native-like slope of initial rise does not affect the phonological/categorical shape of the intonation contours.

2.3. Separation of Phonology and Phonetics in L2 Intonation Development

The results suggest that the phonological characteristics of the target language is acquired before its phonetic characteristics. For example, the advanced learner (JAF) and the intermediate learner (JIM) are phonologically native-like, but phonetically not native-like. It is reasonable to analyze this result in a framework in which the acquisition of speech proceeds along two related but distinct paths. First, L2 learners internalize the phonological structures of L2 intonation, while they still implement these structures by using L1 phonetic habits. Only in a second time, learners master L2 phonetic implementation rules.

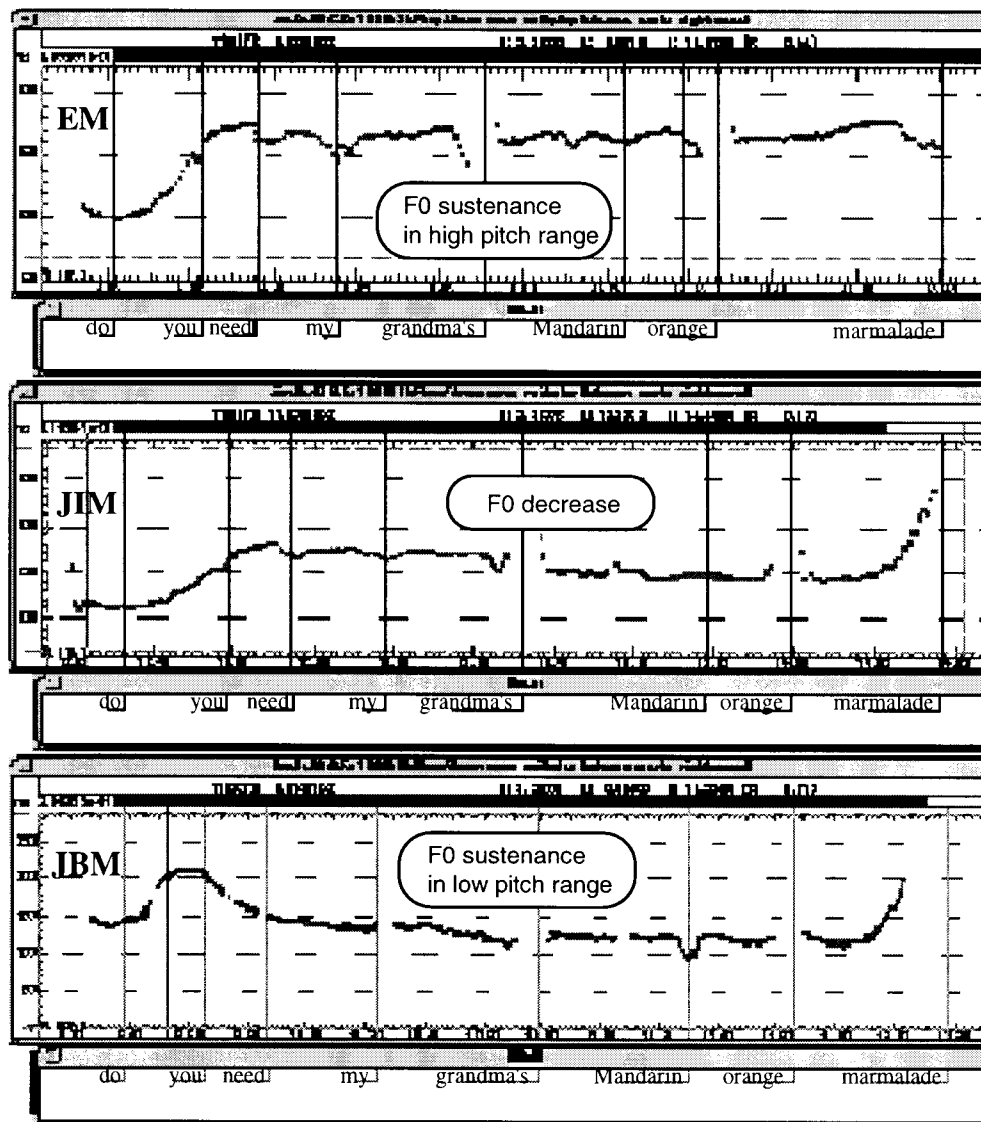


Figure 4. Pitch tracks of EM, JIM and JBM

4. CONCLUSION

The analysis of the experimental data confirms the validity of Ueyama & Jun's claim [6]: the categorical patterns of L2 intonation are acquired before native-like continuous intonational streams. This supports models in which the speech cognitive system is split into a phonological component (characterized by categorical units) and a phonetic component (implementing the phonological units as a continuous articulatory/acoustic stream).

5. REFERENCES

- [1] P. A. Keating, "Phonetic representations in a generative grammar", *Journal of Phonetics*, Vol. 18, pp. 231-334, 1990.
- [2] P. A. Keating, "The window model of coarticulation: articulatory evidence", *Laboratory Phonology I*. Cambridge: Cambridge University Press.
- [3] J. B. Pierrehumbert, *The Phonology and Phonetics of English Intonation*, Ph.D. dissertation, MIT, 1980 [published 1987 by IULC, Bloomington: Indiana University Linguistics Club.]
- [4] M. E. Beckman & J. B. Pierrehumbert, "Intonational structure in Japanese and English", *Phonology Yearbook* Vol. 3, pp. 255-309, 1986.
- [5] D. R. Ladd, "An introduction to intonational phonology", *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, pp. 231-334, Cambridge University Press, 1992.
- [6] M. Ueyama & S-A Jun, "Focus realization in Japanese English and Korean English intonation", *Japanese/Korean Linguistics* Vol. 7, CSLI/Stanford University Press, in press.
- [7] J. B. Pierrehumbert & M. E. Beckman, *Japanese Tone Structure*, Cambridge, MA: MIT Press, 1988.
- [8] K. Maekawa, "Is there 'dephrasing' of the accentual phrase in Japanese?", *Ohio State University Working Papers in Linguistics: Papers from the Linguistic Laboratory*, Vol. 44, pp. 146-165, 1994.