

# PROSODIC FEATURES OF ENGLISH YES-NO QUESTIONS PRODUCED BY JAPANESE SPEAKERS

*Akiyo Joto<sup>a</sup> & Nobuo Yuzawa<sup>b</sup>*

<sup>a</sup>Prefectural University of Hiroshima, Japan; <sup>b</sup>Utsunomiya University, Japan  
joto@pu-hiroshima.ac.jp; yuzawa@cc.utsunomiya-u.ac.jp

## ABSTRACT

This paper examined the characteristics of the English rising tone in yes-no questions produced by native speakers of Japanese according to the different levels of the intelligibility of intonation. Five native English speakers evaluated the intelligibility for three kinds of yes-no questions read by 10 Japanese speakers: the questions were those with the nucleus placed in the sentence-initial, sentence-medial, or sentence-final position. The F0 and the duration of the Japanese speakers' (JS) productions were measured and compared with those of native English speakers' (NS) ones to elucidate acoustic features.

It was found that the intelligibility was lower in the JS productions with the sentence-initial and the sentence-medial nucleus. The acoustic analysis showed that the overall range of the major pitch change was smaller in the JS productions than in the NS ones. The intelligibility was lower in the JS productions with the sentence-initial or sentence-medial nucleus when the magnitude was smaller, and it tended to be higher in controlled JS productions with the sentence-final nucleus when the nucleus had longer duration and was correctly identified. Lower intelligibility was also attributed to the wrong placement of the nucleus and the inappropriate use of nuclear tones. The rise in pitch on the sentence-final word suggested both negative and positive L1 transfer in JS productions.

**Keywords:** English prosody, nucleus, intonation, accent, Japanese speakers

## 1. INTRODUCTION

The English language nowadays plays a vital part as an international language in a socially and linguistically globalized context. Jenkins [4] set up the Lingua Franca Core (LFC) to provide a pedagogical core of phonological intelligibility for speakers of English as an international language, and listed 'nuclear stress production and placement' as one of the items of the LFC. The question then arises as to how intelligible English

nuclear tones produced by JS are to NS's ears. What is the difference between intelligible and unintelligible production?

The objective of this paper is to investigate the characteristics of the English rising tone in yes-no questions produced by JS according to the different levels of intelligibility. Two experiments were conducted to examine the issue: one involved perceiving the English nucleus produced by JS and evaluating the intelligibility of the intonation of the JS productions, and the other was concerned with an acoustical analysis of JS and NS productions.

## 2. EXPERIMENT 1

In this experiment, JS productions were evaluated by NS to investigate their intelligibility.

### 2.1. Materials

The stimulus corpus for Experiment 1 comprised 60 English productions by 10 Japanese speakers (5 females and 5 males). All the productions were context-free yes-no English questions with the nucleus placed in the initial, medial, and final position. These nuclei form 'broad focus' Cruttenden [1] in a sentence. The sentences were as follows with the nucleus in capitals:

- S1. SEEN anything of him?
- S2. Has he aGREED to it?
- S3. Would you like some TEA?

The sentences were taken from Wells [6]. They were read aloud by each JS, first, in their own way (uncontrolled productions; hence, UP), and then, according to the stipulated nucleus placement and nuclear tones (controlled productions; hence, CP). They were digitally recorded, randomly ordered, and recorded onto CDs.

#### 2.1.1. Speakers

All ten Japanese speakers were English learners at the university level in Japan. Their ages ranged from 18 to 21. Two of them had experience of living abroad for more than 4 years and received secondary education overseas by means of the

English language. All the others received standard English education at secondary schools.

Prior to the recordings, each speaker was asked to circle the most accented word in each question on a reading list.

## 2.2. Evaluators

Five native English speakers participated in the perception experiment. They were all students temporarily studying at a university in Japan, who were 19 or 20 years of age. They had lived in Japan for two or three months. Two of them were American, and the rest, British.

## 2.3. Procedure

Each evaluator was given evaluation sheets and a CD containing the JS productions. First they listened to each of the 60 productions on a CD, then on a given evaluation sheet where each sentence was printed they circled the word they heard most accented, i.e., the focused word. If they could not identify any, they were to write a cross (x). Then they were asked to rate the intelligibility of the intonation of a whole sentence on a five-point scale, with 1 as the lowest (very poor) and 5 as the highest (excellent). They wrote reasons for rating, the implications of the speaker's intonation, or comments, if any.

As for CP, two sheets—(1) and (2)—were assigned to each production; one side of sheet (1) was folded to ensure that the evaluators could not see the sentence with the given nucleus written on sheet (2). Evaluators were asked to circle the word they heard most accented on sheet (1) first. They opened the folded sheet (1) and rated the intelligibility of the JS intonation on sheet (2), on which the word was circled in a sentence which a JS intended to place the strongest accent (the nucleus) on.

## 2.4. Results

The mean of the five rating numbers for the intonation of each sentence was calculated for each JS. Table 1 shows the mean of intelligibility of each sentence, the number of JS speakers evaluated at each intelligibility level, and the percentages of the evaluation cases where the nucleus was correctly perceived in both UP and CP.

As shown in Table 1, the intelligibility of intonation was highest in both CP and UP with the sentence-final nucleus (S3), whereas it was lowest in UP with the sentence-medial nucleus (S2) and in CP with the sentence-initial nucleus (S1). It was

more difficult for JS to produce the English nucleus placed in the beginning or middle of a yes-no question than that in the final position.

**Table 1:** The mean of the intelligibility of intonation, the perception of the nucleus (%) and the number of speakers at each intelligibility level in the controlled and uncontrolled productions by Japanese speakers.

sent. No.	intel. levels of intonation	uncontrolled productions			controlled productions				
		mean of intel. level	number of JS	percep. of nucleus (%)	mean of nucleus percep. (%)	mean of intel. level	number of JS	percep. of nucleus (%)	mean of nucleus percep.
1	1.0-1.9		0	0	28.0	2.88	1	60.0	48.0
	2.0-2.9		5	28.0			3	33.3	
	3.0-3.9	3.14	4	30.0			5	52.0	
	4.0-4.9		1	20.0			1	60.0	
	5		0	0			0	0	
2	1.0-1.9		1	40.0	58.0	3.50	0	0	72.0
	2.0-2.9		2	70.0			2	70.0	
	3.0-3.9	3.04	6	56.7			5	76.0	
	4.0-4.9		1	60.0			3	66.7	
	5		0	0			0	0	
3	1.0-1.9		0	0	84.0	3.72	0	0	90.0
	2.0-2.9		0	0			2	70.0	
	3.0-3.9	3.68	7	82.9			4	90.0	
	4.0-4.9		3	86.7			3	100	
	5		0	0			1	100	

In the nuclear placement questionnaire, only two out of ten JS circled 'seen' as the most accented word in S1: six JS selected 'anything,' and two, 'him'. In S2, seven JS circled 'agreed' and in S3 nine chose 'tea.' This could cause the lower intelligibility of S1 and S2 in UP.

The percentages of the nucleus perception were raised in CP, and therefore JS were able to locate the nucleus correctly more often when it was given. However, the intelligibility level did not rise in CP with the sentence-initial nucleus.

Table 1 indicates that the intelligibility and the nucleus perception varied across the JS. The prosodic characteristics at different intelligibility levels from an acoustic perspective will be described in the next section.

## 3. EXPERIMENT 2

The acoustic analysis was carried out for the JS and NS productions in order to discover acoustic characteristics of the JS prosody according to the different intelligibility levels in comparison with those of the NS prosody.

### 3.1. Materials

The JS productions were the same as those used in Experiment 1. The English materials were the same sentences as the JS productions, which were produced by four native English speakers (2 males and 2 females). They consisted of three Americans and one Canadian.

### 3.2. Methods

The lowest F0 value and the highest one at the end of a sentence with rising intonation, or the lowest one at the end with falling intonation were measured for each NS and JS production. As for S3, the lowest and highest F0 values in the nucleus, i.e., 'tea,' were measured. Based on these values, the frequency distance between the highest and the lowest F0 in semitone (cf. Hart, et al. [3]) was calculated in order to obtain the magnitude of the distance. We also measured the duration of the given nucleus in each production, and calculated its percentage in relation to the sentence duration.

Lehiste [5] presents the intrinsic pitch of English vowels, but we did not take it into consideration in this study because the vowels in the JS productions were not always identical with English ones.

### 3.3. Results

Tables 2 and 3 show the results of the acoustic analysis of the NS and JS productions. The number in the parentheses indicates the number of productions.

**Table 2:** The mean of the frequency distances in semitones, the duration, the nuclear tones, and the onset of the major pitch change of the native English speakers' productions.

produc	sent. No.	nuclear tone	onset of major pitch change	ST (mean)	duration (%)
UP	1	Rise (4)	seen (1), anything (3)	10.1	26.7
		Fall-Rise (1)	agreed (1)	7.0	35.4
	2	Rise (3)	agreed (3)	12.1	35.5
		Fall-Rise (1)	agreed (1)	7.0	35.4
	3	Rise (3)	tea (3)	9.9	33.4
		Fall-Rise (1)	tea (1)	7.1	42.9
CP	1	Rise (4)	seen (2), anything (2)	8.3	33.2
		Fall-Rise (1)	agreed (1)	7.1	40.5
	2	Rise (3)	agreed (3)	11.8	41.7
		Fall-Rise (1)	agreed (1)	7.1	40.5
	3	Rise (3)	tea (3)	12.7	35.6
		Fall-Rise (1)	tea (1)	8.4	42.8

As shown in Tables 2 and 3, overall, the frequency distances were smaller in the JS productions than in the NS ones, except for CP of S1 (the mean of ST for the rise: 7.5 for UP and 8.8 for CP of JS, and 10.7 for UP and 10.9 for CP of NS). It indicates that the Japanese speakers used narrower pitch range in English speech productions. In UP in S2 and S3 and in CP in S1 and S2, the frequency distances were bigger in the JS productions with the higher intelligibility level (4.0-5.0) than in those with the lower level (3.9 and under) when the productions had a rising tone.

**Table 3:** The mean of the frequency distances in semitones, the duration, the nuclear tones, and the onset of the major pitch change of the Japanese speakers' productions at the different intelligibility levels of intonation.

produc	sent. No.	intel. level	nuclear tone	onset of major pitch change	ST (mean)	duration (%)	
UP	1	1.0-2.9	Fall (1)	him (1)	-5.9	27.7	
			Fall-Rise(1)	him (1)	3.2	26.9	
		Rise(3)	anything (1), him (2)	10.8	32.2		
		3.0-3.9	Rise (4)	anything (3), him(1)	8.4	29.8	
		4.0-5.0	Rise (1)	him (1)	8.5	28.4	
	2	1.0-2.9	Fall (1)	he (1)	-12.4	20.2	
			Rise (2)	agreed (2)	6.3	33.0	
		3.0-3.9	Rise (6)	agreed (3), it (3)	5.8	25.3	
		4.0-5.0	Rise (1)	agreed (1)	12.9	35.7	
		3	3.0-3.9	Rise (7)	like (2), tea (5)	5.8	28.3
	4.0-5.0		Rise (3)	tea (3)	8.4	29.6	
	CP	1	1.0-2.9	Fall (1)	him (1)	-6.4	37.3
Fall-Rise(1)				anything (1)	14.9	24.6	
Rise (2)			anything(1), him(1)	9.9	31.8		
3.0-3.9			Rise (5)	anything (3), him(2)	9.0	29.9	
4.0-5.0			Rise (1)	anything (1)	11.4	32.7	
2		1.0-2.9	Rise (2)	agreed (1), it (1)	7.1	32.5	
			Rise (5)	agreed (3), it (2)	7.9	31.0	
		4.0-5.0	Rise (3)	agreed (3)	11.5	29.2	
		3	1.0-2.9	Rise (2)	like (2)	6.1	28.7
			3.0-3.9	Rise (4)	tea (4)	9.6	29.6
4.0-5.0		Rise (4)	tea (4)	7.3	35.2		

The duration of the given nucleus in S3 was shorter in the JS productions than in the NS ones in both UP and CP. As for the NS productions, the duration of the given nucleus was longer in CP than in UP. The clear tendency of the duration and the frequency distance was not seen between UP and CP. No clear differences in the duration were found across the JS productions at the different intelligibility levels.

Table 3 presented a falling tone or a fall-rise tone in the JS productions with low intelligibility of intonation of UP in S1 and S2 and of CP in S1, where a rising tone is normally used. One English speaker also used a fall-rise tone in S2 and S3 in both UP and CP, which had different implications.

## 4. DISCUSSION

The correlation was measured between the intelligibility levels of intonation and the frequency distances, the duration, and nucleus identification, and between nucleus identification and the frequency distances and the duration in order to examine the relationships between them. The stronger correlation was seen between the intelligibility and the frequency distances in UP in S2 ( $r=0.54$ ), in CP in S1 ( $r=0.45$ ) and in S2 ( $r=0.43$ ), between the intelligibility and the duration in CP in S3 ( $r=0.39$ ), and also between the intelligibility and nucleus identification in CP in

S3 ( $r=0.50$ ). It indicates that the intelligibility tended to be higher when the magnitude of the pitch range was greater between the lowest F0 and the highest F0 at the end of S2 in UP, and that it was lower when the magnitude was smaller in S1 and S2 in CP. Therefore, we could say that the JS productions with the sentence-initial or the sentence-medial nucleus tended to have lower intelligibility when they were pronounced with the smaller range in pitch between the lowest and the highest at the end of a sentence. This was manifested in the result shown in Table 1 that S1 and S2 had the lower intelligibility than S3.

From the above correlation of intelligibility with nucleus identification and with the duration, we found that the intelligibility of CP in S3 tended to be higher when the duration of the nucleus was longer and the nucleus was correctly identified.

Table 3 indicates that the major pitch change started at the word 'like' in both UP and CP with lower intelligibility whereas the productions with higher intelligibility had the onset of the major pitch change on the word 'tea' which was the nucleus. The acoustic analysis of pitch contour showed that the F0 was higher on 'some' than on 'like' in the NS productions with lower intelligibility. The intelligibility and the identification of the nucleus in S3 were not affected by the pitch height or change on the syllables before 'some.' Therefore, lower pitch on 'some' than on 'like' was a crucial point for the identification of the nucleus and for the intelligibility of this sentence. These cases showed the wrong placement of the nucleus where the JS put the nucleus on 'like,' not 'tea.'

Close investigation of the acoustical features of the JS productions and Table 3 indicated that the rise in pitch often occurred on the last word of a sentence like 'him' in S1 or 'it' in S2 with lower intelligibility. This phenomenon was not seen in the NS productions. This might have been affected by the intonation of questions in the Japanese language where the rise occurs on the last syllable of a question marker 'ka' at the end of questions. The degree of difficulty, in the framework of behaviorist theories, depends on the extent to which the L2 pattern is similar to or different from the L1 pattern. When the two are identical, learning could take place easily through 'positive transfer' of the L1 pattern, and when they are different, learning difficulty and errors occur as a result of 'negative transfer' Ellis [2]. It is considered that negative L1 transfer occurred in S1

and S2 with lower intelligibility whereas positive L1 transfer was seen in S3 which had the highest intelligibility as shown in Table 1. Therefore, it would be easier for Japanese speakers to produce yes-no questions with the nucleus at the end.

## 5. CONCLUSIONS

We examined the prosodic features of the English rising intonation of yes-no questions produced by Japanese speakers according to the different intelligibility levels of intonation. It was found that the intelligibility differed across the positions of the nucleus in a sentence, and that it was lower when the nucleus was placed in the initial or middle position. The acoustical analysis showed that the magnitude of pitch range was related with the differences in intelligibility across the positions of the nucleus.

Lower intelligibility of intonation was attributed to the wrong placement of the nucleus and the inappropriate use of nuclear tones: the fall and/or the rise-fall were seen in the JS productions with lower intelligibility in S1 and S2. The rise on a sentence-final word suggested that both negative and positive L1 transfer would have occurred in the JS productions.

## 6. ACKNOWLEDGEMENTS

This study was supported by a Grant-in-Aid for Scientific Research promoted by the Japan Society for the Promotion of Science (Grant no. 22520505).

## 7. REFERENCES

- [1] Cruttenden, A. 1986. *Intonation*. Cambridge: Cambridge University Press.
- [2] Ellis, R. 2008. *The Study of Second Language Acquisition* (2<sup>nd</sup> ed.). Oxford: Oxford University Press.
- [3] 't Hart, J., Collier, R., Cohen, A. 1990. *A Perceptual Study of Intonation: An Experimental-phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.
- [4] Jenkins, J. 2000. *The Phonology of English as an International Language*. Oxford: Oxford University Press.
- [5] Lehiste, I. 1970. *Suprasegmentals*. Cambridge, Mass.: the MIT Press.
- [6] Wells, J.C. 2006. *English Intonation: An Introduction*. Cambridge: Cambridge University Press.