

Effect of Interactive Visual Feedback on the Improvement of English Intonation of Japanese EFL Learners

Masaki Taniguchi

0. Introduction and Purpose

This research was dedicated to helping to improve the teaching and learning of English intonation (prosody) for Japanese EFL learners. It attempted to evaluate the effectiveness of the use of real time interactive visual feedback on the learners' approximation of their fundamental frequency contours to native speakers'.

1. Material and Method

Four kinds of material were used: the well-known passage, *The North Wind and the Sun* (a standard text in phonetic research) (See *The Principles of the International Phonetic Association* 1949) without tone marks, a group of short phrases with six kinds of nuclear tone marks, a dialogue without tone marks and a dialogue with tone marks. They were all presented in written form and shown to the learners in orthography (not in phonetic transcription). The tone marks used were from the O'Connor and Arnold system of 1973, except that we did not follow them in the height differentiation of the degree sign apparently intended by them to indicate stressed but unaccented syllables.

Twelve learners were used as informants and were divided into two groups, a control and a test group, with six members in each. The learners were all female Japanese college EFL students who attended the Summer Course in English Phonetics (SCEP) at the Department of Phonetics and Linguistics, University College London, 1996. SCEP 1996 took place from 29 July to 9 August. It consisted of ten sessions spread over the two weeks. All the students attended the lectures together and were divided into classes of ten or so for practical lessons. Two hour-long practical lessons were given each day, one on segmental aspects of pronunciation and the other mainly on prosodic features.

Ten of the informants, five in each group, had specialized in English linguistics, and therefore had some knowledge of phonetics, including intonation. Two (one in each group) had specialized in law and had no knowledge of phonetics before attending SCEP. About half of them in each group had a pronunciation of Japanese close to the standard, and the other half a pronunciation with no pitch accent distinction. They were asked to read the material aloud twice, once before SCEP and once again after SCEP. Before SCEP, they were given no instructions other than to read the material aloud.

Stereo DAT recordings were made of their speech (microphone recording) and simultaneous larynx output (Lx). (Lx information is obtained by monitoring vocal fold vibrations with externally placed laryngograph electrodes.)

During the SCEP, both groups attended all the required lectures and practical lessons, including the ones on intonation. The difference between the control group and the test group was that the control group did not use Laryngograph-based visual feedback at all, while the test group used it regularly in their intonation lessons. Each group had a tutor, a native speaker of English, for their practical lesson on intonation as well as for the one on segmental pronunciation. Both groups regularly used the same teaching material out of a

textbook, not including the test material, except for a brief look at it with some explanation of the tone marks on the test dialogues and phrases. We examined how far each group improved by attending SCEP, comparing their performance after SCEP with their performance before SCEP. We also examined how the use of an interactive visual display affects their improvement by comparing the test group with the control group.

The material used was as follows (M = Material):

M1. The North Wind and the Sun: A Passage Presented Without Tone Marks

The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him, and at last the North Wind gave up the attempt. Then the Sun shone out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

M2. Dialogue Presented Without Tone Marks

- A: Good morning.
 B: Good morning.
 A: Your passport, please.
 B: Here you are.
 A: Is this your first visit to Britain?
 B: No, it's my second visit.
 A: How long are you here for?
 B: I'm here for a month.
 A: For sightseeing?
 B: Well, partly for sightseeing.
 A: What's the main purpose of your visit?
 B: To attend a summer course.
 A: What kind of course is it?
 B: It's a course in English phonetics.
 A: Where?
 B: At University College London.
 A: All right. Enjoy your stay.
 B: Thank you.

M3. Short Phrases Presented With Tone Marks

- | | | |
|----------------|----------------|----------------|
| `No | `No | ^No |
| ^No | ^No | ^No |
| `No one | `No one | ^No one |
| ^No one | ^No one | ^No one |
| `No one has | `No one has | ^No one has |
| ^No one has | ^No one has | ^No one has |
| `No one has it | `No one has it | ^No one has it |
| ^No one has it | ^No one has it | ^No one has it |

M4. Dialogue Presented With Tone Marks

- A: How `old are you?
 B: 'How old do you `think I am?
 A: I don't `know.
 B: You 'don't ,know? 'How old do I ,look?
 A: >Well, | I've 'no i ,dea.
 B: ^No?
 A: Seven ,teen?
 B: ~No.
 A: How old `are you?
 B: 'Twenty-`two.
 A: ~Twenty-~two.
 B: ^Yes.
 A: You ,look so `young.
 B: I ~don't look °young.
 A: You ~do.

2.1. Basic Idea behind the Use of Interactive Visual Feedback

The Laryngograph Processor is a device which operates with a personal computer using a suite of speech processing programs (PCLX), which analyses laryngeal excitation waveforms (Lx), and speech fundamental frequency (Fx). One of its characteristic features is that it can display the Fx contours of the speaker in real time on the screen of the computer attached to it. It can also display two persons' Fx contours on the same computer screen, for instance, the model speaker's in the top half and the learner's in the bottom.

We must note here that Fx and pitch are not the same. The traces on the screen are Fx contours, which are, however, the major physical correlate of pitch. Pitch is an auditory perceptual category which depends on fundamental frequency variation: we perceive a rising pitch, for example, when the Fx values change from lower to higher.

What this study aimed at was to find out whether we could more effectively help the Japanese learner to produce the contours reasonably matching the model speaker in the appropriate places in an utterance. When we try to copy a visual fundamental frequency contour, we are not trying to change the absolute fundamental frequency values, but only to imitate the general shape, which is called "pattern matching," and this is the basic idea behind the use of interactive visual feedback. The user is imitating the overall pattern and not aiming to produce the exact fundamental frequency values; this is called "normalisation" and is at the heart of speech perception. Therefore this paper will not show any Fx values although it would be possible to do so, but will only discuss the general shapes of the Fx contours.

3. Findings**3.1. Number of Unnatural Accents/Stresses and Word Boundaries**

The number of unnatural word boundaries was counted only in M1 (a passage without

tone marks). The code "a" indicates the number of syllables with unnatural/unusual accents/stresses, "b" the number of unnatural word boundaries. Tables 1a to 4 below show the numbers of syllables with unnatural/unusual accents/stresses and word boundaries. Table 1a shows those of the accents/stresses in M1, Table 1b shows those of word boundaries in M1, Table 2 shows those of the accents/stresses in M2, Table 3 shows those of the accents/stresses in M3, and Table 4 shows those of the accents/stresses in M4. Figures 1a to 4 correspond to Tables 1a to 4.

A, B, C, D, E and F are control group informants; a-f is their mean. G, H, I, J, K and L are test group informants; g-l is their mean. "BC" means "Before SCEP" and "AC" "After SCEP."

Table 1a: Number of Syllables with Unnatural Accents/Stresses in M1

	A	B	C	D	E	F	a-f	G	H	I	J	K	L	g-l
BC	10	15	26	20	13	26	18.3	10	12	16	9	9	19	12.5
AC	13	18	29	12	10	26	18	9	7	7	10	9	19	10.2

Table 1b: Number of Unnatural Word Boundaries in M1

	A	B	C	D	E	F	a-f	G	H	I	J	K	L	g-l
BC	4	15	18	14	11	19	13.5	4	8	11	13	12	26	12.3
AC	2	12	25	13	5	22	13.2	7	7	3	7	11	24	9.8

Table 2: Number of Syllables with Unnatural Accents/Stresses in M2

	A	B	C	D	E	F	a-f	G	H	I	J	K	L	g-l
BC	10	21	26	14	16	19	17.7	26	15	19	17	17	16	18.3
AC	13	24	23	12	14	16	17	5	6	13	7	9	16	9.3

Table 3: Number of Syllables with Unnatural Accents/Stresses in M3

	A	B	C	D	E	F	a-f	G	H	I	J	K	L	g-l
BC	16	14	16	12	9	17	14	19	9	16	16	11	20	15.2
AC	2	4	5	5	2	7	4.2	4	2	8	1	0	3	3

Table 4: Number of Syllables with Unnatural Accents/Stresses in M4

	A	B	C	D	E	F	a-f	G	H	I	J	K	L	g-l
BC	18	25	26	22	17	18	21	30	30	18	18	14	21	21.8
AC	17	17	10	16	12	17	14.8	10	10	12	13	9	16	11.7

Figure 1a

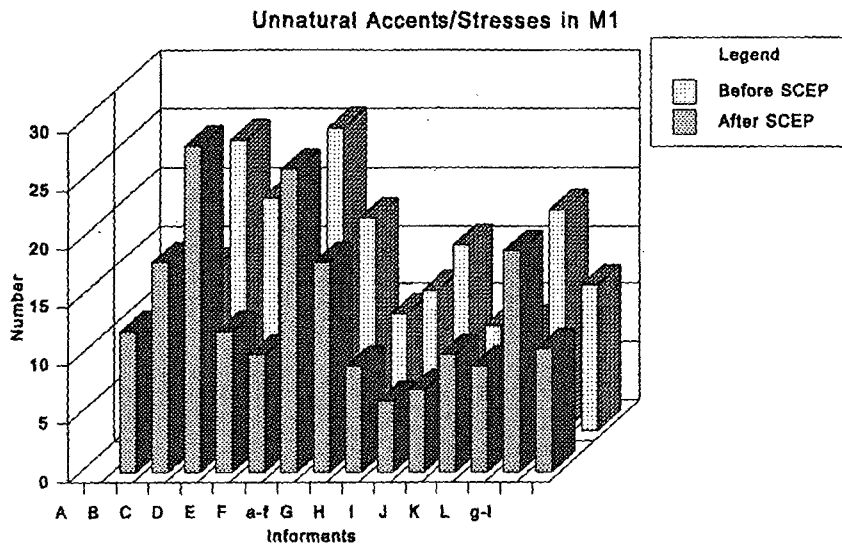


Figure 1b

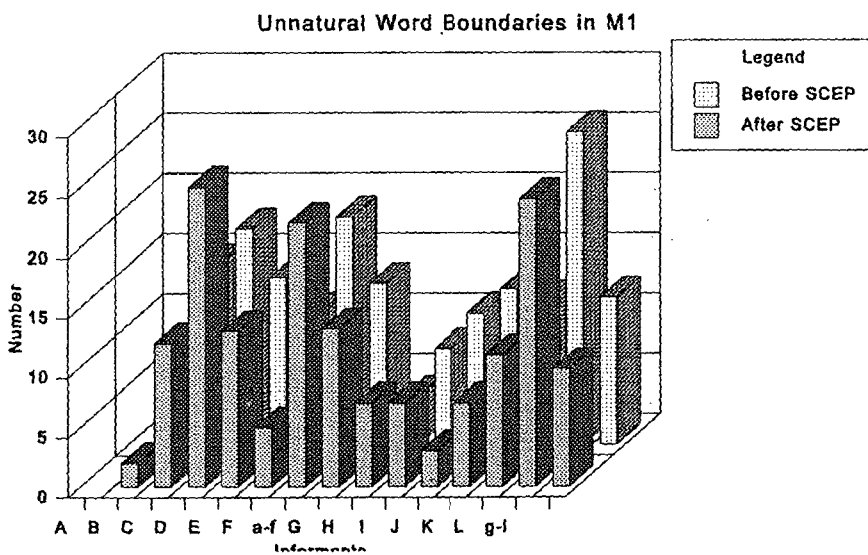


Figure 2

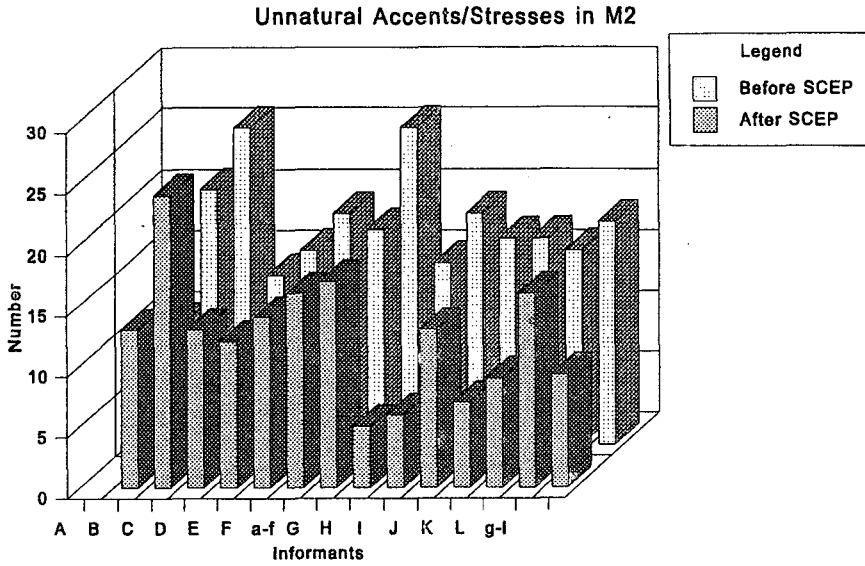


Figure 3

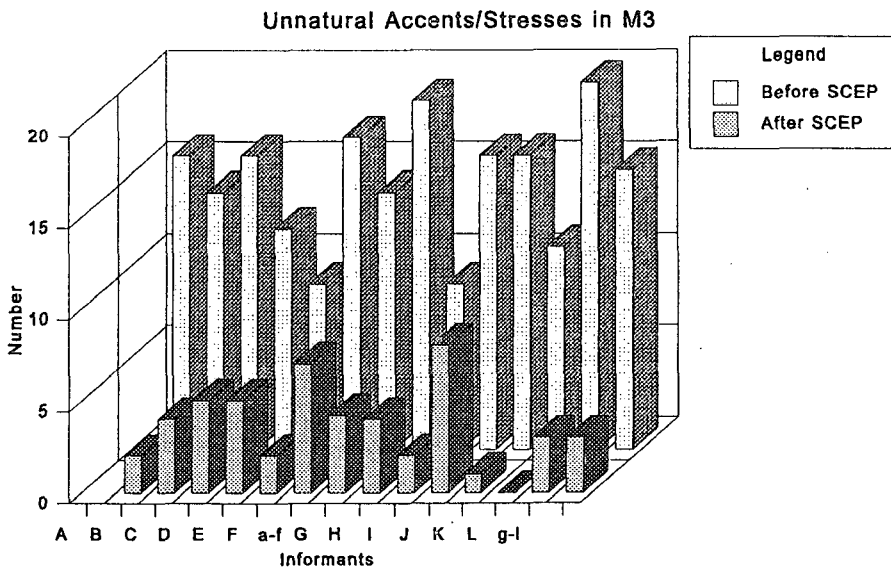
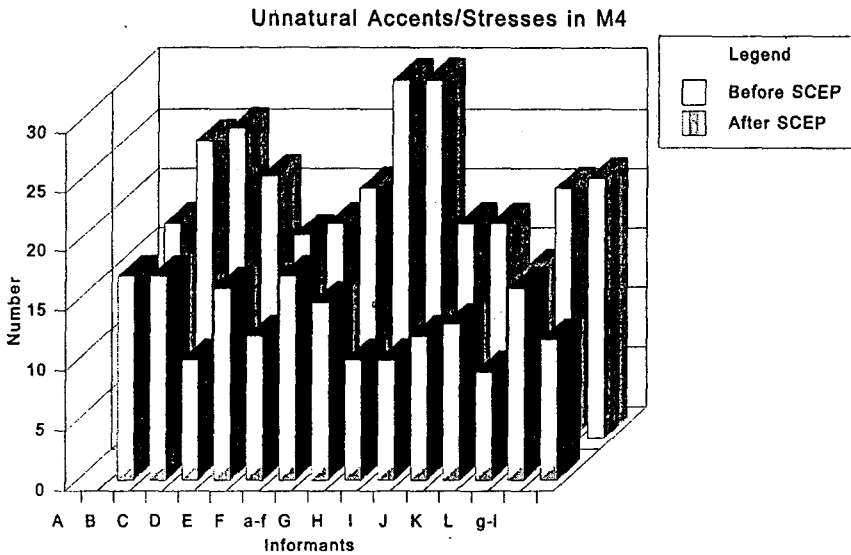


Figure 4



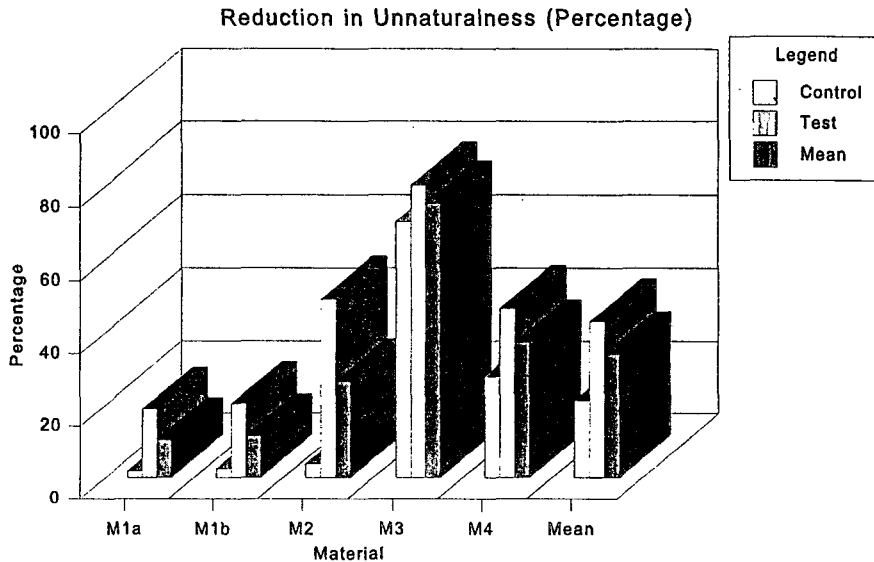
3.2. Improvement Percentages

Table 5 and Figure 5 show in percentages how much the control group and the test group reduced the numbers of their syllables with unnatural accents/stresses and unnatural word boundaries. "M1," "M2," etc. indicate the same as above, and so are "a" and "b." "Control" indicates "control group" and "Test" indicates "test group."

Table 5: Percentage of Reduction in the Number of Syllables with Unnatural Accents/Syllables and Unnatural Word Boundaries

	M1a	M1b	M2	M3	M4	Mean
Control	1.8	2.4	3.8	70.2	27.9	21.2
Test	18.8	20.3	49.1	80.2	46.7	43.0
Mean	10.3	11.4	26.5	75.2	37.3	34.1

Figure 5



3.3. Results

The results that we can see in Table 5 and Figure 5 tell us several things. Although both the control and test groups improved in general during the SCEP, but there was a clear difference between the two groups as well as between the kinds of material used. (M1: a passage without tone marks; M2: a dialogue without tone marks; M3: short phrases with tone marks; and M4: a dialogue with tone marks)

First, the test group surpassed the control group in each type of material used. This implies that the use of interactive visual display with the Laryngograph greatly helped the test group students to improve their naturalness in English intonation.

Second, both groups displayed greater improvement in the material *with* tone marks than in the material *without* tone marks. This implies that the use of tone marks helped the learners to improve their intonation.

Third, the difference between the two groups was greater in the material *without* tone marks than in the material *with* tone marks.

Fourth, the control group improved only in the material *with* tone marks, whereas the test group improved not only in the material *with* tone marks but even in the material *without* tone marks. This implies that, even in the material *without* tone marks, the use of interactive visual feedback with the Laryngograph greatly helped the test group students to improve their naturalness in English intonation.

The results with the test group indicates that it was very difficult for students to grasp the natural intonation if the material did not carry tone marks and if they were not provided with interactive visual feedback.

Fifth, the highest improvement was obtained in M3 with both groups (higher in the test group than in the control group), which consisted of short phrases of increasing length with tone marks. This implies that such short phrases as "No," "No one," "No one has," and "No one has it" were very effective in teaching English intonation to Japanese learners although the location of the nuclear tone did not vary.

Overall, the highest achievement in the learners' English intonation can be expected when both interactive visual display and tone marks are used. To repeat one of the above findings, which was really very striking: in the material *without* tone marks, the control group made no improvement, but the test group, with the help of the Laryngograph interactive visual feedback, did make substantial improvement.

3.4. Some Characteristics of Japanese Learners' English

The following characteristics of Japanese learners' English were observed in this investigation.

3.4.1. The Japanese learners' English intonation, etc. seemed to be influenced by the pitch patterns of Japanese. For instance, in words and compounds like "passport" and "summer course," even though the primary stress is placed on the first syllable in each of them, their counterparts (loan words) in Japanese have the following patterns: [pas(w)po[^]oto] (pitch accent: L(H)HLL (L= Low, H=High)(^ = pitch accent falls at this mora boundary), and [samaa + ko[^]os(w)] (pitch accent: LHHHL(L)). Thus many of the informants in the present investigation failed to produce the appropriate accentuation and gave primary stress to a later syllable. For example, they used the pattern, 'pass_port instead of `passport, and 'summer_course instead of `summer course.

3.4.2. The Japanese learners' intonation tended generally to be narrow in pitch range and it was often difficult to discern what pitch pattern was being employed. This seemed to be also an influence of a general characteristic of Japanese intonation.

3.4.3. The Japanese learners' intonation exhibited what we might term 'over-chunking,' that is, dividing into too many separate word groups, hence unsuitably many nuclear tones. When teaching Japanese learners to make their English intonation more dynamic, we may have to be careful and teach them to do so only at the appropriate locations and not all through their utterances. For instance, Student F2 in this investigation used the following pattern in M1: ..._one | `who first | suc_ceeded | in `making | the `traveller | `take his | `cloak off should be | `considered | `stronger than the | `other.|

3.4.4. More important than the above even, very often Japanese learners seemed to fail to place the nuclear tone at the appropriate location, by failing to see the contrast between the key words, such as "first" and "second" (in M2). For instance, Student E (both E1 and E2) failed to place the nuclear tone on "second" but used the following pattern:

A: Is 'this your °first °visit to 'Britain? |

B: `No, | it's my 'se°cond `visit. |

They also seemed to fail to notice that some words are old information and therefore best be deaccented, for instance, "here" in "I'm here for a month." in response to "How long are you here for?" (in M2); "sightseeing" in "Well, partly for sightseeing." in response to "For sightseeing?" (in M2); and "young" in "I don't look young." in response to "You look so

young." (in M4). Student A2 used the following pattern:

A: How 'long are you `here for? |

B: I'm `here for a `month. |

Student B2 used the following pattern:

A: For ,sightseeing? |

B: >Well, | >partly for | `sightseeing.

Student D2 used the following pattern:

A: You \look so `young.

B: I \don't look `young.

Since the same word is used in each of these cases, it is advisable to deaccent it when it is used the second time at a point very close to its first occurrence.

3.4.5. Many Japanese learners failed to deaccent in what might seemed to be fairly obvious cases of old information, let alone where it was not that obvious, for instance, "purpose of your visit" in "What's the main purpose of your visit?" in response to "Well, partly for sightseeing." (in M2). Since "sightseeing" is a purpose of visit, it is advisable to deaccent "purpose of your visit" in "What's the main purpose of your visit?" However, many of the informants in this investigation missed this point. For instance, Student C1 used the following pattern:

B: \Well, | >partly | >for | 'sight`seeing. |

A: 'What's the >main | `purpose of your | `visit? |

3.4.6. Japanese learners also seemed to have difficulty in choosing or using the appropriate nuclear tones. For instance, many of the informants in this investigation used a High Fall tone on the first word in "Here you are" (in M2) and had no subsequent Rise with no ingratiating voice quality and/or temper features. It may be a possible pattern in some situations, but certainly not a preferred one in this situation where a tourist is talking to an immigration officer at an airport. It may sound too blunt. For instance, Student B (both B1 and B2) used the following pattern:

A: Your `passport, ,please. |

B: `Here you are. |

The appropriate nuclear tone would be a Fall-plus-Rise. This is what usually happens with native speakers, so it is good for EFL learners to follow these general tendencies. To help them learn more natural and fluent intonation with appropriate shape and location of nuclear tones, the use of interactive visual feedback seemed to be highly effective.

4. Conclusion

This investigation fortunately enabled us to reaffirm our confidence in the effectiveness of the interactive visual feedback provided by the Laryngograph Processor in helping Japanese EFL learners improve their English intonation. We saw that there was a great difference in improvement between the group of learners who had the advantage of being exposed to interactive visual feedback for an hour every day in the two-week course and the group of learners who did not. We also found that the use of tone marks helped the learners a great deal, but an important finding was that if no tone marks were provided, it was extremely difficult for the learners to improve without any interactive visual feedback. With the use of interactive visual feedback, the learners were able to improve even in material without tone marks.