

FOR IMPROVEMENT OF ENGLISH INTONATION LEARNING SYSTEM

Hiroshi Suzuki\*, Shigeru Kiritani and Hiroshi Imagawa

1. INTRODUCTION

The high performance of recent personal computers has made various kinds of real-time processing of speech signals possible. We have devised a system, with a personal computer, which analyzes and displays fundamental frequencies of speech sounds nearly in real-time), and we have attempted to use it as a training device of English stress and intonation for Japanese learners of English. A set of model utterances of English words and sentences for training are put into the system in advance, and a specified utterance is played back with its pitch and intensity patterns displayed on the screen. The learner listens to the model and puts in his own utterance through the microphone. The pitch and intensity patterns of his utterance are drawn superposed onto the patterns of the model so that any discrepancy can be readily noticed, and the learner sees what part of his utterances should be improved. The display of the learner's patterns are in a different color from that of the model.

2. HOW THE SYSTEM WORKS

The initial menu on the screen enables the learner to choose a word or a sentence he wants to learn. Once he chooses one, its pitch and intensity pattern appears on the screen in a few seconds. With that on the screen he can do any of the following by pushing a key on the keyboard.

1) Listen to the model utterance.

The speech can be reproduced as many times as he wishes.

2) Put in his own utterance mimicking the model and have the pitch and intensity pattern of it on the screen.

Once he puts in his utterance through the microphone, the speech signals are processed and the visual patterns appear in a matter of a few seconds. The intensity pattern appears a little above that of the model to avoid merging of the two intensity patterns.

3) Normalize the learner's patterns.

The learner's overall pitch is usually different from that of the speaker of the model, and the speech tempo may also be different. In order to make it easier to compare the learner's patterns with those of the model, the system has a normalizing function in terms of pitch and tempo. The pitch normalization is based on the average frequency of the utterance.

4) Listen to his own utterance.

He can readily compare his own utterance with the model

---

\*English Department, College of Arts and Sciences, University of Tokyo

utterance, reproducing one after the other.

5) Have the list of words and sentences for learning on the screen.

This enables the learner to choose another word or sentence to learn.

It often helps the learner if a pair of sentences contrastive to each other such as affirmative versus negative are shown on the same screen. In order to realize this, our system has two independent displays of the patterns described above: one on the upper half of the screen and the other on the lower half. A sample screen is shown in Fig. 1.

### 3. TEACHING MATERIALS

In order to test the feasibility of the system, some words and sentences were prepared and read by a native speaker of American English. They were then put into the system, and each part of the patterns was labelled with orthographical representations. Besides, vertical lines were drawn to indicate the boundaries of words in the case of sentences for learning intonation, and the place where the accent falls in the case of sentences for learning sentence stress.

There are three categories in the materials prepared: namely, intonation, word stress, and sentence stress pattern. The words and sentences used were as follow.

#### [INTONATION]

No. 1	This is my car.	(indicative)
No. 2	Come and see me tomorrow.	(imperative)
No. 3	What are you looking for?	(wh-question)
No. 4	How did you do it?	(wh-question)
No. 5	Are you going to stay here?	(yes-no question)
No. 6	Is this for me?	(yes-no question)
No. 7	I have visited, Korea, China, and India.	(additive enumeration)
No. 8	Which would you like, tea or coffee?	(alternative question)
No. 9	Would you like tea or coffee?	(yea-no question)
No.10	He is our new teacher, isn't he?	(tag-question expressin certainty)
No.11	You're hungry, aren't you?	(tag-question expressin uncertainty)
No.12	It's my mother who did it.	(emphatic)
No.13	Is it George who needs it?	(emphatic)

#### [WORD STRESS]

No. 1	influence	(first syllable)
No. 2	realize	( " )
No. 3	territory	( " )
No. 4	official	(second syllable)
No. 5	economic	(third syllable)
No. 6	intellectual	( " )
No. 7	physiology	( " )

[SENTENCE STRESS]

No. 1	I want to meet him.	(- / - / -)
No. 2	He mustn't leave her.	(- / - / -)
No. 3	He started to talk to me.	(- / - - / - -)
No. 4	He came on a bicycle.	(- / - - / - -)
No. 5	I think it will be fine.	(- / - - - /)
No. 6	I'll see him in a week.	(- / - - - /)
No. 7	I think that he wants us to go.	(- / - - / - - //)
No. 8	It isn't the same as before.	(- / - - / - - //)
No. 9	I think it was an excellent idea.	(- / - - - / - - - //)
No.10	We finished it the day before he came.	(- / - - - / - - - //)
No.11	The woman has fallen and broken her leg.	(- / - - / - - / - - /
No.12	We haven't got time to arrange for it now.	(- / - - / - - / - - /

4. TRIAL USE

A number of Japanese college students have used the system, and it was found that it helps them learn English stress and intonation. Fig. 2 and Fig. 3 are two samples of the patterns obtained with one of those students. His first mimicry after listening to the model utterance several times is shown in Fig. 2-A and Fig. 3-A, and his final utterance after learning with the audio and visual display is shown in Fig. 2-B and Fig. 3-B. The solid lines in the pitch patterns indicate the model, and the dotted lines the learner's performance.

In Fig. 2 improvement is clearly seen with the contrast of rising tone for "tea" and falling tone for "coffee." The first half of the sentence shows marked improvement. With Fig. 3, some improvement is observed in the isochronic rhythmic pattern.

Some of the students learned by themselves and others received instructional assistance. Generally speaking, the students learned faster with instructional assistance, partly because the assistance helped them identify the correspondence between the part of their utterances and the part on the curve of the screen.

5. REMARKS

It was ascertained that the system is a useful tool for learning stress, intonation, and rhythmic patterns of English, thanks to semi-real-time response and the overlay function.

We found that the most important problem to tackle is to determine the acceptable range of curve deviation. The students tried repeatedly to make their curves identical with that of the model. Although there were still some discrepancies between theirs and the model, their speech sounded quite satisfactory in most cases. Research into acceptability level and its visual representation should be needed to make this system really feasible.

REFERENCES

1) Imagawa, H., and S. Kiritani: High-speed speech analysis system using a personal computer and a DSP, Ann. Bull. RILP, 22, 1989.

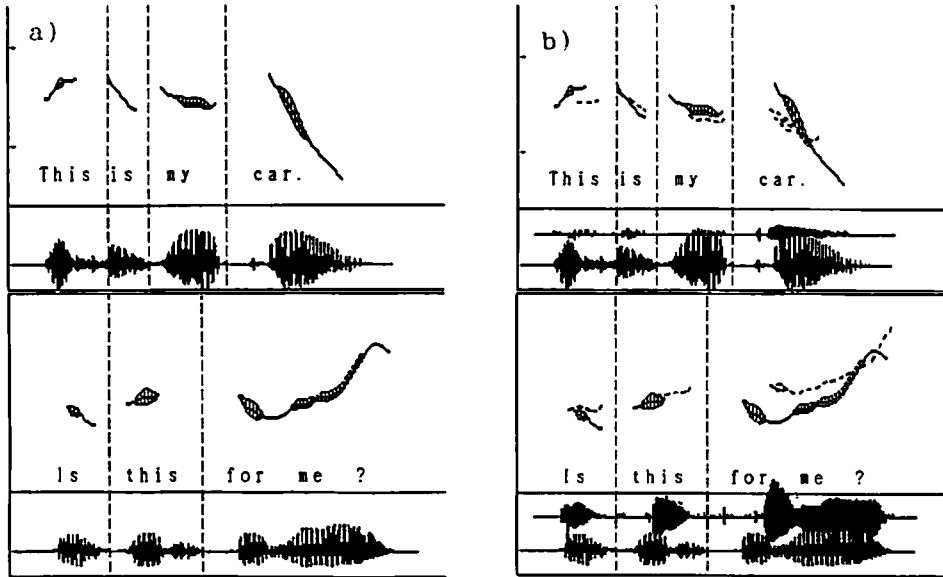


Figure 1 (a) Patterns of the model utterance.  
 (b) Student's patterns superposed onto the patterns of the model.

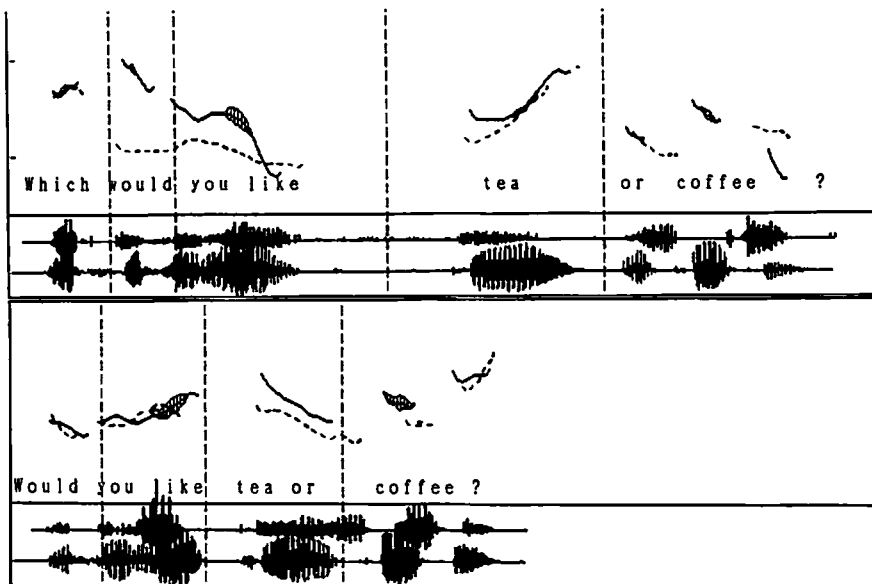


Figure 2 (a) Student's first attempt.

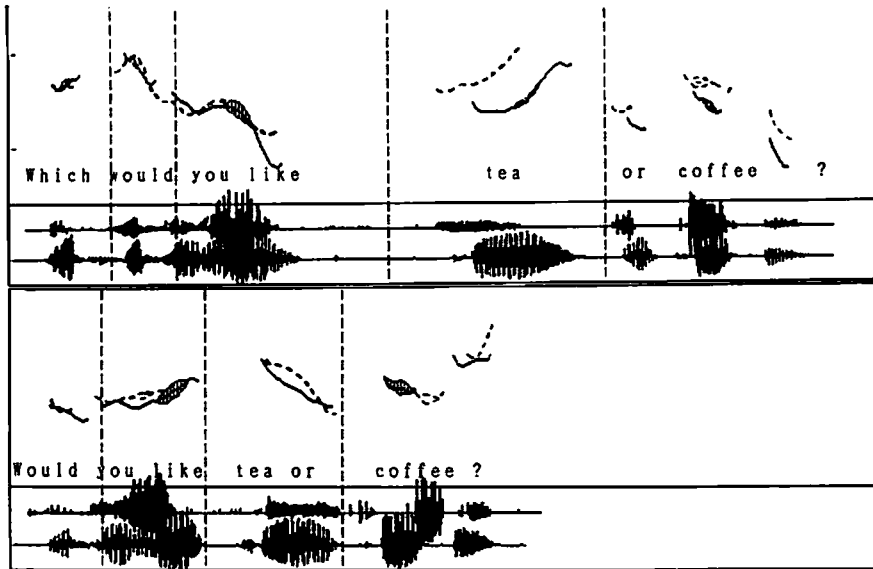


Figure 2 (b) Student's final attempt.

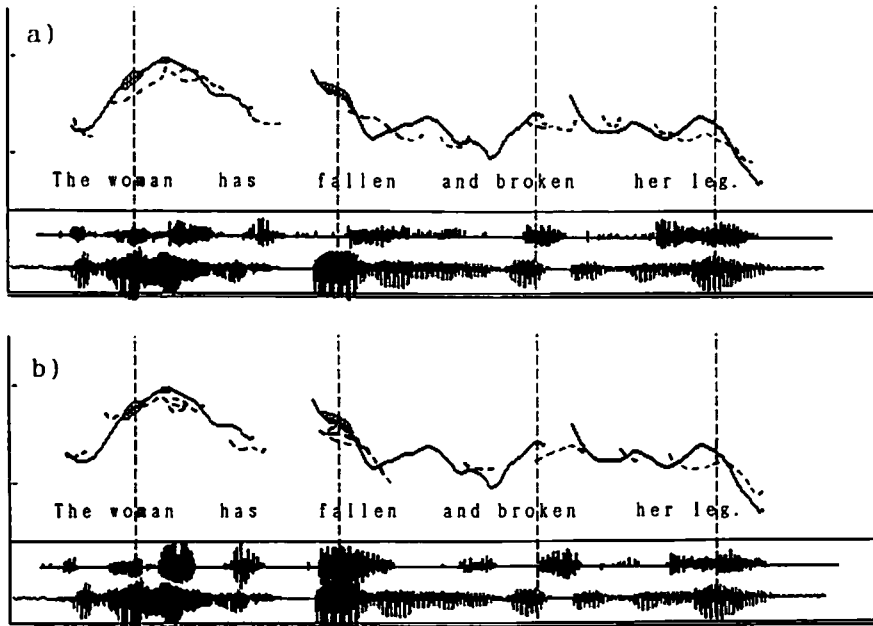


Figure 3 (a) Student's first attempt.  
 (b) Student's final attempt.