Perspectives

Teaching Suprasegmentals to Japanese Learners of English Through Electronic Visual Feedback

Janet Anderson-Hsieh
Iowa State University

After discussing the importance of suprasegmentals (stress, rhythm, and intonation) in communication and the effectiveness of electronic visual feedback (EVF) for their instruction, this paper explains the specific ways that EVF can be used to teach suprasegmentals to Japanese learners of English. Contrasts between English and Japanese suprasegmentals are described and the difficulties that Japanese learners demonstrate with English stress and intonation resulting from these contrasts are discussed. Procedures and exercises for teaching stress and intonation are explained, and examples of EVF graphs are presented. Finally, the advantages and disadvantages of using EVF in Japan are identified, and it is concluded that EVF is a powerful tool for raising Japanese learners' awareness of the important role of suprasegmentals in communication.

本論は、コミュニケーションにおける超文節的単位（ストレス、リズム、イントネーション）の重要性と、これらの項目の指導時のコンピュータを使った視覚的フィードバックの有効性を論じた後、日本人英語学習者に英語の超文節的単位を教える際の、コンピュータを使った視覚的フィードバックの具体的な使用法を解説する。英語と日本語の超文節的単位の違いが記述され、それらの相違が原因となる、日本人学習者が英語のストレスとイントネーションを習得する際の困難点が論じられる。ストレスとイントネーションを教える手順と練習方法が解説され、日本においてコンピュータを使った視覚的フィードバックを使用することの利点と欠点が指摘される。結論として、コミュニケーションにおける超文節的単位の重要性を日本人学習者に気づかせるためには、コンピュータを使った視覚的フィードバックは強力な手段であるという主張がなされる。

Suprasegmentals (stress, rhythm, and intonation) are critical in communication because they provide the framework for utterances. They highlight the information the speaker regards as important, and they convey the speaker's intentions and emotional states (Bolinger, 1986; Cruttenden, 1986; Kenworthy, 1987). It has therefore been argued

(Dalton & Seidlhofer, 1994; Dickerson, 1989; Gilbert, 1984; Kenworthy, 1987; Wong, 1987) that ESL learners need to pay particular attention to suprasegmentals, perhaps even more than to segmentals. However, when listening to spoken discourse, learners so often focus on individual lexical items that they tend to ignore the overriding melody and rhythm of utterances, and because they do not readily perceive these suprasegmentals, they tend to have difficulty producing them.

Electronic equipment exists which addresses this problem by providing visual feedback on suprasegmentals in real time. As a student speaks into a microphone connected to the equipment, pitch and intensity—the physical correlates of suprasegmentals—are extracted from the speech signal and displayed on a video screen while the student is speaking. A dual display presents the student's pitch and intensity contours on one half of the screen and those of a native speaker model on the other half, allowing comparisons to be made in timing, intonation, and stress. Several types of computer software and hardware with electronic visual feedback (EVF) capabilities for displaying suprasegmentals are now on the market. Among the most frequently used are the Visi-Pitch (1987) and the IBM Speech Viewer (Undated).

Research has shown that EVF is an effective tool for teaching intonation to ESL learners, and in the last decade, EVF has been used in ESL programs throughout the United States for teaching various aspects of suprasegmentals (Anderson-Hsieh, 1992; 1994; Chun, 1989; Lane, Mitchell, Molholt, Pennington, Perdreau, Cessaris & Fischer, 1987; Molholt, 1988; Pennington, 1989). The purpose of this paper is to describe how EVF can be used to address the problems that Japanese learners of English have with English stress and intonation. First, suprasegmental differences between Japanese and English will be discussed. Then an approach to teaching word-level stress using EVF will be described. This will be followed by a description of how to use EVF to teach sentence stress and intonation. Finally, some general guidelines will be presented, and recommendations will be made for using EVF in Japan.

Suprasegmental Differences Between Japanese and English

Important differences exist between English and Japanese in the way stress is realized in speech. In English, stressed syllables are highly marked, generally showing greater pitch movement, higher intensity (loudness), and longer duration than unstressed syllables. Here, pitch and duration are the most important features in the perception of stress (Fry, 1957). In contrast, stress, or accent, in Japanese is less marked

because only pitch is used. Duration is used in Japanese for lexical distinctions such as that between obasan, "aunt," and obaasan , "grandmother," but it is not a feature of stress, as it is in English (Vance, 1987).

Also, differences exist between Japanese and English in the way pitch is used to mark stress in words. In polysyllabic English words spoken in isolation using final falling intonation, stress is marked on a non-initial stressed syllable by a jump up in pitch from the unstressed to the stressed syllable followed immediately by a fall. This pattern is illustrated in the following polysyllabic English words:

(1) re · ception (2) ther · mometer (3) re · ceive

Stress, or accent, is realized quite differently in Japanese (Vance, 1987). Several patterns of non-initial stress in Japanese words are presented below. The letter "L" above the syllable represents a low pitch or tone, and the letter "H" a high pitch or tone.

(4)	(5)	(6)	(7)
L H H	L H H L	L H H H	L H L
/taka ra/	/kaga ribi/	/kamiso ri/	/ko koro/
"treasure"	"bonfire"	"razor"	"heart"

(Vance, 1987, p.. 79-80)

The first observation which can be made from these examples is that while a stressed syllable is always made with a high tone, unstressed syllables can be realized as either high or low tones, except for initial unstressed syllables and unstressed syllables occurring after the stressed syllable. In these two cases, stress is always realized as a low tone. Furthermore, if the last syllable of the word is stressed—assuming that the word is utterance-final—there is no fall from a high tone to a low tone because there is no following unstressed syllable to which the pitch can step down or fall. Only in the pattern (7) above is a stressed syllable (a high tone), immediately preceded and followed by an unstressed syllable (both realized as low tones). It is also important to note that in Japanese, some words such as *sakana* for "fish" are not stressed at all. Thus, it can be seen that some very important differences exist in the way stress is realized through pitch in English and Japanese.

The differences between English and Japanese in duration and pitch discussed above are at least in part responsible for some of the difficulties that Japanese speakers have in pronouncing stressed syllables correctly in English. The beginning or intermediate Japanese learner tends not to sufficiently lengthen stressed syllables and may not always show a sufficient rise in pitch on the stressed syllable followed by a sufficient fall. The stressed syllables will therefore tend not to stand out as much

as they will for a native speaker of English, and this can adversely affect intelligibility.

Teaching Word Level Stress

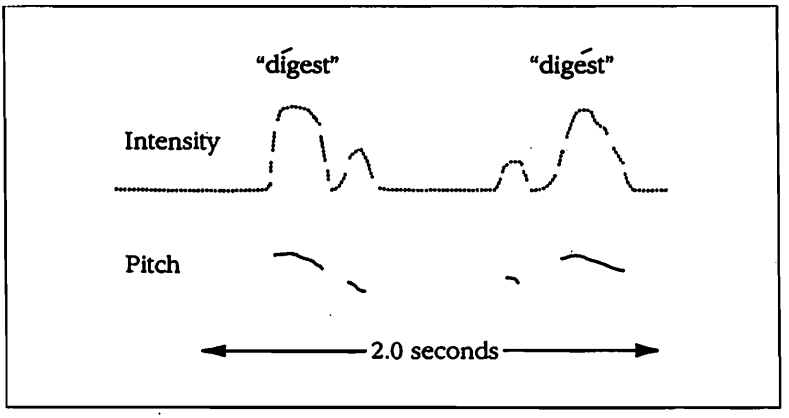
When correcting for stress problems at the word level, the teacher can first work with pairs of words which are identical except for their stress patterns. Such pairs can be found in noun and verb forms such as those in the following list:

Noun	*Verb*
digest	digest
project	project
record	record
subject	subject

Using the EVF, the teacher speaks into the microphone, pronouncing one pair of words at a time. The dual display is used, with the teacher's pitch and intensity contours appearing in the upper half of the display and the student's appearing later in the lower half.

To demonstrate the features of English stress on the EVF visual display, the noun and verb forms of "digest" as spoken by a native speaker of English are presented in the EVF graph in Figure 1. The equipment that was used was the Visi-Pitch. The intensity (loudness) contour appears in the upper half of the graph and the pitch contour appears in the lower half. The time, which is represented on the horizontal axis, is

Figure 1.



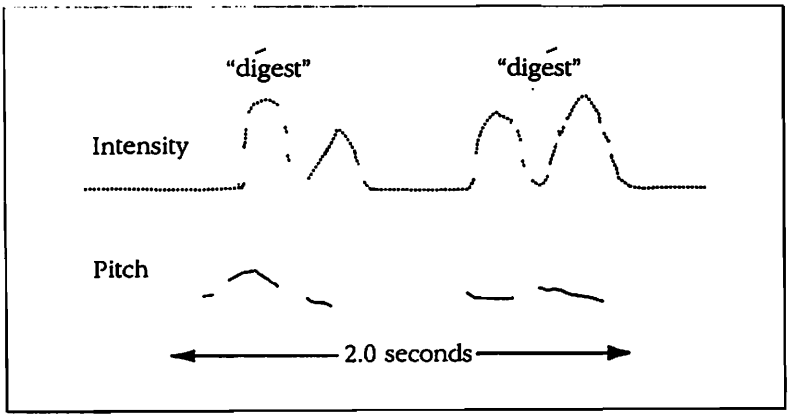
2.0 seconds. When comparing the same syllable in stressed and unstressed position, it can be seen that the syllable in the stressed position is more intense, longer in duration, and higher in pitch than in the unstressed position. This is consistent with what was described in the discussion of English stress given above.

After explaining these features of stressed syllables in English using the visual display as an example, the teacher then asks the student to speak the same two words into the microphone. The student's pitch and intensity contours appear on the screen just below the native speaker model. The teacher then asks the student to comment on any differences s/he sees between the native speaker model and the display of the student's own speech.

The EVF graph of a Japanese speaker (TOEFL 397) pronouncing the noun and verb forms of "digest" appears in Figure 2. When comparing this graph to the native speaker graph in Figure 1, it can be seen that, unlike the native speaker's forms, the student's forms show little difference in loudness and duration between the stressed and unstressed forms of the same syllable. Also, although the pitch pattern for the noun is acceptable, the pitch pattern for the verb shows a failure to jump up sufficiently on the stressed syllable.

After discussing the visual display with the student, the teacher then encourages the student to repeat the words again, striving for greater differentiation in length between stressed and unstressed syllables and higher pitch on the stressed syllables. It is important for the teacher to explain that an exact replication is not necessary, for even a native

Figure 2.



speaker's speech can vary somewhat from one repetition to the next. In fact, the teacher might even demonstrate this variation on the visual display and perhaps use other models of English to show that variation also exists among native speakers from different regions.

What is important, however, are the relative differences between the student's pronunciation and the teacher's. Thus, the student must strive for longer stressed syllables, shorter unstressed ones, and a higher pitch on the stressed syllables, although the pitch need not show exactly the same degree of height as the teacher's. The student's goal should be to consistently show patterns that are closer to the native speaker model on several successive attempts before moving on to the other forms on the list. In the beginning, the teacher will have to make the decision as to whether the student has come close enough and should guide the student accordingly.

Teaching Sentence-Level Stress and Intonation

After the student shows some progress in using pitch, duration, and intensity correctly to mark stress in isolated words, the teacher can then move on to teaching stress and intonation at the thought-group level. A thought group is a phrase, a clause, or a sentence bordered by pauses on both sides. In thought groups, one stressed syllable stands out "heads and shoulders" above all the other stressed syllables in the sentence. This syllable, which is said to carry the main sentence stress (or accent), is marked by a more pronounced movement in pitch than is found on the other stressed syllables in the thought group.

In unmarked or neutral English sentences, the main sentence stress falls on the stressed syllable of the last content word in the sentence. This is illustrated in the sentence below in which the word receiving main stress is presented in upper case letters.

John said he would LEAVE.

In marked sentences, however, the accent can fall just about anywhere else, and the shift in stress is usually due to contrast, contradiction, or a shift in focus. In the marked sentence presented below, the sentence stress has shifted to the first word, and the sentence can be interpreted as contradiction.

JOHN [not Bill] said he would leave.

In teaching sentence-level stress, the teacher should present sentence pairs, such as the above, for practice and the procedures and suggestions for modeling and practice described under "word-level stress" should be used.

The EVF graph in Figure 3, in which a 4.5 second time display is used, shows how the previous two sentences, as spoken by a native speaker of English, differ in their pitch, intensity, and duration patterns. The unmarked first sentence appears on the left-hand side and the marked second sentence appears on the right-hand side. When comparing the two sentences, especially notable in the graph is the higher pitch and much longer duration of "John" in the marked sentence and "leave" in the unmarked sentence. The final intonation pattern for both sentences is falling intonation, typically found in statements and information questions.

The Japanese learner's versions of the same two sentences are presented in Figure 4. When comparing the two sentences, it can be seen that their pitch and intensity patterns show little difference, even though the student reported that he thought he had succeeded in producing them differently. The variation in syllable duration so noticeable in the native speaker's utterances in Figure 3 is much less noticeable in the Japanese learner's syllable duration in Figure 4. Also, there is a failure to jump up sufficiently in pitch for the stressed word "leave" in the first sentence. Such "syllable-timed" rhythm and failure to highlight stressed syllables sufficiently using pitch can render the student's speech more difficult to understand. No doubt the intended message of the second sentence, contradiction or correction, would have been lost on the listener. The student must obviously learn to both lengthen sufficiently the syllable receiving the main stress and jump up in pitch sufficiently at the same time.

After the student identifies these problems and can clearly see the difference between the native speaker model and his or her own speech,

Figure 3.

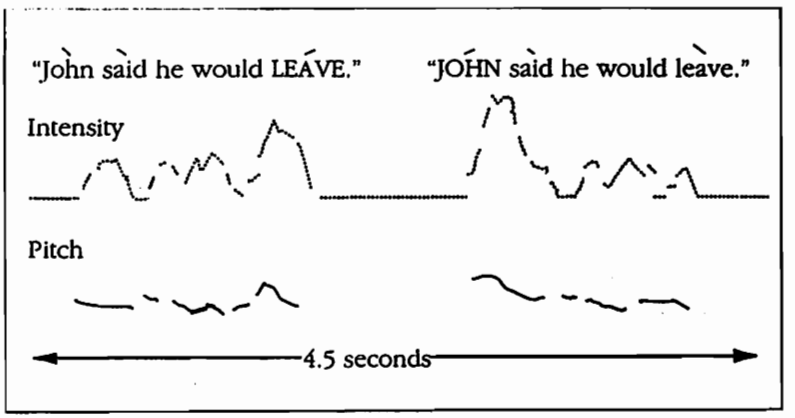
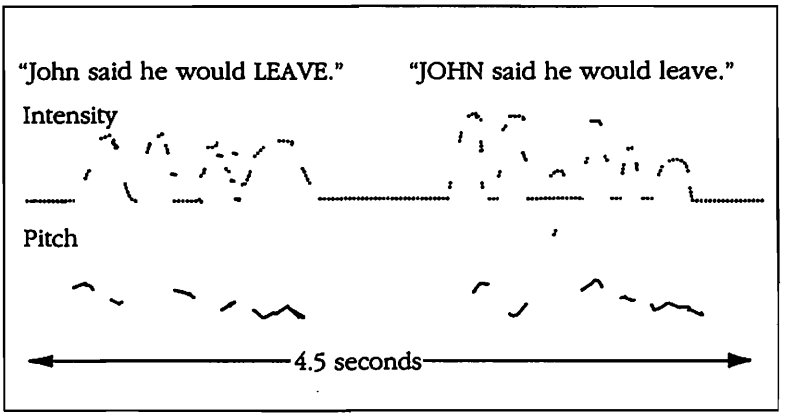


Figure 4.



s/he should then begin repeating the sentences using EVF until more native-like patterns begin to appear. Then sets of sentences such as the following can be used, in which the syllable receiving the main stress occurs in several different positions instead of in only two.

John bought a new CAR	(neutral sentence stress)
John bought a NEW car	(not a used one)
John BOUGHT a new car	(he didn't rent one)
JOHN bought a new car	(not Harry)

Dialogues such as the following, which clearly demonstrate the important link between meaning and intonation in English, can also be used quite effectively with EVF for practicing intonation.

A: I want some shoes.
B: What kind of shoes?
A: The beautiful kind.
B. Black or brown?
A. Neither. I'm tired of black and brown.
B. I want red shoes. Shiny red shoes. (Gilbert, 1993, p. 88)

In all of the above sentences, the basic intonation patterns are the same—they all show final falling intonation. Other intonation patterns, such as final rising intonation for expressing uncertainty and asking questions and the fall-short rise pattern for non-final intonation should also be introduced fairly early using both "canned dialogues" and ex-

cerpts from prerecorded samples of spontaneous speech.

Another type of problem which can be addressed through EVF is the Japanese learner's tendency to insert extra vowels in consonant clusters and after final consonants. This is due to the "katakana effect," the insertion of vowels in such environments in foreign borrowings. Although extra vowels did not appear in the speech of the intermediate Japanese learner which was illustrated in Figures 2 and 4 above, such tendencies are well known and have been documented in Anderson-Hsieh, Riney and Koehler (1994). In that study, it was shown that vowel epenthesis (insertion) occurred 12.7% of the time in predicted phonetic environments during a reading task for intermediate Japanese learners of English and only 2.8% of the time for advanced learners. Such extra vowels are very easy to identify on the EVF display, where they appear as additional peaks on the intensity display. To correct this tendency, the student practices suppressing the extra vowel until the visual display no longer shows extra peaks of intensity.

Some General Guidelines for Using EVF

When using EVF, the teacher can work with students individually or in very small groups in which the students take turns using the equipment and observe and learn from others in the group while waiting their turn. It is very important to teach students to read the display quite early so that they can take responsibility for their own learning, discovering for themselves the differences between the representations of their own speech and those of the native speaker model. If a printer is available, hard copies can be made of some of the EVF graphs, and the students can take them home as reminders of what they have worked on with EVF.

In addition, the teacher must realize the limitations of EVF—that it is basically a tool for repetitive practice (Morley, 1994)—and that for students to learn really well, they must also practice outside of class and begin to transfer what they are learning to communicative situations in their lives. Homework assignments should be given that require the students to monitor their speech and use the patterns they are working on with EVF communicatively.

Furthermore, in Japan and other EFL situations where native-speaker teachers may not always be available, the teacher can use tape-recorded speech of English speakers with EVF if necessary. Most types of EVF equipment should allow input from a tape recorder or other external source.

Conclusions

This report has presented some suggestions for using EVF effectively with Japanese learners of English to improve their stress and intonation patterns. Learners who have used EVF often report that they find it helpful for improving suprasegmentals, an element of speech production which they did not fully understand until they used EVF. They often say that they did not know how their stress and intonation patterns differed from native speech until they saw their utterances displayed on a screen and were able to compare them with those of a native speaker model in real time. Other advantages of EVF are an increase in student motivation and the convenience of EVF for explaining students' errors. The major disadvantage of EVF is that it may be too costly for some schools to afford and it is not convenient for use in large classes except perhaps for demonstration. However, that EVF is a powerful tool for raising Japanese language learners' awareness of suprasegmentals is indisputable.

Acknowledgments

I am indebted to one of the reviewer's helpful suggestions which I have incorporated into the paper. I am also indebted to the graduate college at Iowa State University for the equipment grant which was used to purchase the EVF feedback equipment used in this paper.

Janet Anderson-Hsieh, a professor in the Department of English at Iowa State University, teaches in the TESL, ESL and Linguistics. Her recent publications on pronunciation teaching and interlanguage phonology have appeared in *Language Learning*, the *TESOL Quarterly*, the *JALT Journal*, *System*, the *CALICO Journal* and *IDEAL*.

References

Anderson-Hsieh, J. (1992). Using electronic visual feedback to teach suprasegmentals. *System, 20*, 51-62.

Anderson-Hsieh, J. (1994). Interpreting visual feedback in computer assisted pronunciation instruction on suprasegmentals. *CALICO Journal, 11*(4), 5-22.

Anderson-Hsieh, J., Riney, T. & Koehler, K. (1994). Connected speech modifications in the English of Japanese ESL learners. *IDEAL, 7*, 31-52.

Bolinger, D. (1986). *Intonation and its parts*. Palo Alto, CA: Stanford University Press.

Chun, D. (1989). Teaching tone and intonation with microcomputers. *CALICO Journal, 6*(3), 21-47.

Cruttenden, A. (1986). *Intonation.* Cambridge: Cambridge University Press.
Dalton, C. & Seidlehofer, B. (1994). *Pronunciation.* Oxford: Oxford University Press.
Dickerson, W. (1989). *Stress in the stream of speech: Teacher's manual.* Urbana: University of Illinois Press.
Fry, D. (1958). Experiments in the perception of stress. *Language and Speech, 1,* 126-152.
Gilbert, J. (1984). *Clear speech.* Cambridge: Cambridge University Press.
Gilbert, J. (1993). *Clear speech, 2nd ed.* Cambridge: Cambridge University Press.
IBM Speech Viewer Personal System/2: A guide to clinical and educational applications. (Undated). IBM.
Kenworthy, J. (1987). *Teaching English pronunciation.* London: Longman.
Molholt, G. (1988). Computer assisted instruction for Chinese speakers. *TESOL Quarterly, 22,* 91-111.
Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25* (3), 481-520.
Lane, L. Mitchell, P. Molholt G., Pennington, M., Perdreau, C., Cessaris, A. & Fischer, L. (1988). Taking the mystery out of the mouth: Visible speech. Paper presented at the 22nd Annual TESOL Convention, Chicago.
Pennington, M. (1989). Applications of computers in the development of speaking and listening proficiency. In M. Pennington. (Ed.), *Teaching English with computers* (pp. 99-121). La Jolla, CA: Athelston Publications.
Vance, T. (1987). *An introduction to Japanese phonology.* Albany, NY: State University of New York Press.
VisiPitch IBM Interface, Model 6097 & 6098: Instruction manual. (1987). Pinebrook, NJ: Kay Elemetrics Corp.
Wong, R. (1987). Teaching pronunciation: Focus on rhythm and intonation. Washington, DC: TESOL Publications.

(Received February 16, 1996; revised April 15, 1996)

English the Whole World Hears

Adopted by 200 universities!

SIMUL ACADEMY

MASTER COURSE

with CNN INTERNATIONAL

Video-Based English Text

- ***Basic Level***
  (TOEIC 400－550)
- ***Intermediate Level***
  (TOEIC 550－730)

- contemporary English as broadcast around the world by CNN
- real people, places and events
- a wide range of topics from popular culture to social and environmental issues and light human interest to hard news
- from CNN programs such as *World News, Newsroom, Business Asia, Travel Guide* and *Showbiz Today*

Course Materials

★ **Student's Workbook**
Price: ¥2,000
12 units (Each unit includes Pre-Viewing Exercises, Video Exercises and Expansion Exercises.)

★ **Video Tape**
Price: ¥9,800

★ **Teacher's Manual**
Price: ¥3,000
Includes Teaching Notes, Annotated Transcripts of all reports, Answer Key and Optional Exercises.

TO ORDER, PLEASE CALL THE SIMUL PRESS SALES DEPARTMENT AT **03-3582-4271**

THE SIMUL PRESS, INC.

KOWA BLDG. No.9 1-8-10, AKASAKA, MINATO-KU, TOKYO 102 JAPAN