



Title	Teaching the Suprasegmental Features of English
Author(s)	大熊, 栄
Citation	明治大学教養論集, 319: 37-53
URL	<a href="http://hdl.handle.net/10291/10119">http://hdl.handle.net/10291/10119</a>
Rights	
Issue Date	1999-03-31
Text version	publisher
Type	Departmental Bulletin Paper
DOI	

<https://m-repo.lib.meiji.ac.jp/>

# Teaching the Suprasegmental Features of English

Sakae Okuma

## Contents

1. Introduction
  - 1.1 Reception and production
  - 1.2 How to understand the stream of speech in English
  - 1.3 Comparison of the suprasegmental features of English with those of Japanese, our first language
  - 1.4 Three principles
  
2. The component parts of the suprasegmental features of English
  - 2.1 Features and systems
  - 2.2 Tone and intonation
    - 2.2.1 (Tone as a part of intonation)
    - 2.2.2 (The form of intonation)
    - 2.2.3 (Levels and contours)
    - 2.2.4 (The functions of intonation)
  - 2.3 Stress and rhythm
    - 2.3.1 (Stress as prominence)
    - 2.3.2 (Rhythm as stress time)
  
3. Stress time and mora time

3.1 Mora

3.2 Rhythm and tone

4. Conclusion

5. References

## 1. Introduction

### 1.1 Reception and production

Teaching pronunciation systematically does not mean teaching phonological theories. From our point of view, the suprasegmental features of English have only two aspects: receptive and productive. In the situation of teaching English as a Foreign Language (EFL) which is my main concern in this paper, the receptive aspect comes first before productive one. This is because, in the EFL situation, we do not use English in our daily life. If we try to, we can surely listen to English through radio and television, but we have few chances to use English on a day-to-day basis.

So the first principle of teaching the suprasegmental features of English systematically in the EFL situation should be to teach the receptive aspect of them first, and the productive one next.

### 1.2 How to understand a stream of speech in English

The component parts of language are phonology, morphology, and syntax, so any stream of speech in English consists of those three parts theoretically. Morphological and syntactical knowledge is not enough to understand the speech. Phonological knowledge and practice are also required.

Phonology consists of two branches: the segmental branch and the

suprasegmental one. The phonological description usually begins with the segmental features of sound and moves on to the suprasegmental features. But, from the practical point of view, we should teach the suprasegmental features first as a principle. This is because teaching the segments like consonants and vowels does not always lead to understanding the whole stream of speech. We should begin with larger elements like stress, rhythm and intonation, because they are the first clues to understand the stream which is another name for chaos to learners.

### **1.3 Comparison of the suprasegmental features of English with those of Japanese, our first language**

Learners of a foreign language tend to carry their first language into the target language, giving the features of the first language to the production of the foreign language. This is called first language interference and the phenomena can be found also in the other areas, that is, morphology and syntax. Pronunciation is especially prone to be affected by the first language and so the Japanese production of English sounds is sometimes called “Japlish.”

The first language interference may be inevitable, but it is necessary to remove it as much as possible, for we need to be better understood when we produce English sounds. Our knowledge of both the target language and the first one may help us to remove the first language interference. In this sense it is necessary to teach the comparative phonology, in our case, between English and Japanese.

It is true that there are lots of problems for us to solve in the branch of segmental features, such as the pronunciation of [l] and [r]. But, again, the comparison of the suprasegmental features should come first as a principle. If we want to know the fundamental difference between the two languages, the

comparison of stress time with mora time is indispensable, however limited our curriculum time of teaching may be.

## **1.4 Three principles**

Thus, there are three principles underlying systematic guidance in teaching the suprasegmental features of English in the EFL situation: the first one is to teach the receptive aspect of those features first, and the productive one next. The second one is to teach the suprasegmental features first, and the segmental features next. The third one is to teach the comparative phonology between English and Japanese, especially the comparison of stress time with mora time, to remove the first language interference as much as possible.

## **2. The component parts of the suprasegmental features of English**

### **2.1 Features and systems**

It is not easy to define what the suprasegmental features of English are. Some literature defines them as length, pitch, and stress, while other literature defines them as tone, intonation and rhythm. Even pause and juncture are sometimes included in the features. We have to consider why this variety happens.

We can find a clue in Gimson (1989; 57), who says “a sound has not only quality, ... but also length, pitch, and a degree of stress. Such features may extend in time beyond the limits of the phoneme and embrace much higher units of the utterance. ... such features are prosodic, or suprasegmental.”

Gimson, thus, think of length, pitch, and stress as the suprasegmental “features” of “a sound”. Clark and Yallop (1995; 330) exclude stress out of those three and put loudness into the group. It is more reasonable, they maintain, to call pitch, loudness and duration (or length) the suprasegmental “features” of a sound, for stress is a result of “systematic exploitation of these features” (ibid.; 330).

The systematic exploitation of the suprasegmental features produces not only stress but also tone, intonation and rhythm. These are “systems”, not the “features of a sound”, though they are usually called the suprasegmentals. They are systems in the sense that they are functional. For example, the meaning of the word “import” differs according to the location of stress; and the rising tone can mean a query. (Tench (1996) uses the term “system” in the same way.)

It is not so easy, again, to distinguish these systems from each other, because they do not function independently of each other. In English, for example, tone is a part of intonation and stress has much to do with rhythm.

From a practical point of view of teaching, the suprasegmental systems of a speech are more important than the features of a sound. So we have to know more about stress, rhythm, tone and intonation.

## **2.2 Tone and intonation**

### **2.2.1 (Tone as a part of intonation)**

Tone and intonation can be converged on one and we can call it intonation.

The term tone is almost synonymous with pitch or pitch patterns in English, and only a limited number of separate tones are recognized such as high, low, falling, rising. The tone as a system is not a mere suprasegmental feature, but functional. If we say “no” with the rising tone, we usually give a

meaning to the tone and the meaning may be a query. In this sense the tone is not simply identical with pitch or pitch patterns, but a part of intonation.

The tone is vitally important in a language like Chinese which is called a tone language. The single syllable [ma] can be four words according to the tones. The tone is “a feature of lexicon” (Cruttenden, 1997; 8). Japanese also uses the tone in the same way.

English does not use the tone in such a way. The tone is closely related to intonation.

### 2.2.2 (The form of intonation)

The study of intonation has two aspects: form and function.

To begin with its form, analysts of English intonation propose to break an utterance into what they call tone-units (Roach 1991; 144, Kreidler, 1989; 156) or intonation-groups (Cruttenden, 1997; 8).

It is not easy to recognize what is called a tone-unit boundary. It is especially difficult to do so in “spontaneous speech (particularly conversation)” (Cruttenden, 1997; 29).

Theoretically, the number of tone-units within an utterance is not fixed, depending upon the speaker’s attitude. Cruttenden (1997, 68) points out that “we can if we wish give every syllable a separate intonation-group”, but he adds that “even for purposes of emphasis this is rare.”

Nevertheless, according to Cruttenden (1997; 30-35) again, the markers of tone-unit (or intonation-group) boundaries are known to us and they are pauses, anacruses (like “and he was” in “I saw John yesterday and he was just off to London”), and the final syllable lengthenings.

The structure of a tone-unit is shown by Roach (1991; 147) as:

(PH) (H) TS (T),

meaning pre-head, head, tonic syllable, tail (brackets indicates optional components). A pre-head can be called an anacrusis in Cruttenden's term.

### 2.2.3 (Levels and contours)

There have been two approaches to the description of *English intonation* in the last fifty years; the levels approach which American analysts have preferred and the contour approach which British analysts have preferred.

The levels approach, whatever assumption it may be based on, proposes four contrasting levels of pitch;

4——high  
 3——high-mid  
 2——low-mid  
 1——low  
 (Nasr, 1997; 43)

These levels are claimed to be relative, but critics against this approach point out that the levels must be absolute, "for otherwise this type of analysis becomes wholly arbitrary" (Cruttenden, 1997; 38). They also point out that the levels can be three, or five.

On the contrary, the contour approach analyzes intonation in terms of rises and falls and proposes such contours as long fall, short fall, long rise, short rise, rise-fall, and fall-rise.

Cruttenden introduces the recent two-level approach. This approach has been developed to avoid some of the problems of the four-level approach.



It seems important for us to know that the pitch levels are relative, not absolute and to practice the changes of pitch relatively.

#### 2.2.4 (The functions of intonation)

Roach (1991; 147) indicates the four traditional functions of intonation; attitudinal, accentual, grammatical, and discourse. The last three are called syntagmatic functions.

From the receptive point of view of learning, the knowledge of the functions of intonation will be of great help for listening comprehension. So we have to teach at least the following generalizations.

The attitudinal function is controversial and we have only a few generalizations about it. For example, a falling contour often suggests finality; and a rising contour may indicate “the speaker’s expectation of some response from the listener”, or “a contradiction of what has been said previously” (Kreidler, 1989; 192–193). A rising contour can also mean encouraging.

A fall-rise contour can mean uncertainty, doubt, or requesting.

A rise-fall contour can mean surprise, or being impressed. (Roach, 1991; 168)

The accentual function concerns with the placement of the tonic syllable. The most common placement is on the last lexical word of the tone-unit. But we can place it on any word for contrastive purposes such as emphasis.

The grammatical function of intonation can help the listener to recognize the grammatical and syntactic structure of an utterance. This function means the placement of boundaries between phrases, clauses or sentences. For example (the mark[/]indicates a boundary),

- (a) Those who sold quickly/made a profit.

(b) Those who sold/quickly made a profit.

(Roach, 1991; 174)

The intonation of question-tags (e.g. “isn’t it”, “can’t it”, etc.) is included in this grammatical function.

The discourse function of intonation is “a comparatively new area of study” (Roach, 1991; 176) and it covers a larger unit than a sentence. There are two main functions; one of them is to focus the listener’s attention on the most important aspect of the message or the highest information content. There is a generalized rule:

“the more predictable a word’s occurrence is in a given context, the lower its information content is” (Roach, 1991; 176).

What is called the intonation subordination is included in the function of “attention focusing”.

Another main intonational discourse function is the regulation of conversational behaviour, that is, “speakers use various prosodic components to indicate to others that they have finished speaking, that another person is expected to speak, that a particular type of response is required, and so on” (Roach, 1991; 178).

Thus, intonation is not so easy to understand, but our conclusion of this section is that it should be taught first among other suprasegmental systems and features. This is because intonation is closely related to a whole utterance and it is the whole utterance, not any parts of it, that confronts us and demands our listening comprehension.

## 2.3 Stress and rhythm

### 2.3.1 (Stress as prominence)

The definition of stress is controversial. Some literature defines it as “the force of breath” (Jones, 1956; 141, Nasr, 1997; 42, etc). But Cruttenden (1997; 13), for example, calls this type of definition “misguided”, since “loudness generally plays a minor role in producing prominences”. He himself defines it simply as “prominence”.

In our view, however, Cruttenden’s definition itself is only one side of the same coin. He defines the stress from the receiver’s side.

It is Pennington (1996; 129) who defines the term stress in two ways, that is, for the speaker and for the hearer. For the speaker (or we would like to say the producer), the stress is “the amount of effort or energy expended in producing a syllable”, and for the hearer (or the receiver), the stress is realized as “perceptual prominence, or strength”.

In the previous section 2.1, we employed Clark and Yallop’s definition of stress as a result of systematic exploitation of the suprasegmental features, that is, pitch, loudness and duration. This definition is obviously for the producer. Clark and Yallop (1995; 340) gives the receiver-sided definition as well by saying, “stress is a conventional label for the overall prominence of certain syllables relative to others within a linguistic system.”

We can summarize that the stress is prominence for the receiver.

### 2.3.2 (Rhythm as stress time)

In English, rhythm as a beat is based on the stressed syllable. This language has fairly a strong tendency that the syllables between stresses are compressed into the same time. So, to talk about English rhythm is to talk

about stress time.

The most concise definition of stress time may be that “English has an isochrony based on stresses” (Cruttenden, 1997; 20). This means that we take a roughly equal amount of time from one stressed syllable to the next.

Two terms proposed by Cruttenden are useful for understanding the structure of stress time. One is a rhythm-group and the other is anacrusis. He uses quite similar terms for explaining the structure of intonation like intonation-groups and anacrusis. This implies that stress time overlaps with intonation in the field of rhythm which is based on stresses. But the greatest difference between them lies in that intonation has four functions while stress time does not.

The rhythm group is “the stretch of utterance from one stressed syllable to the next (including a stressed syllable with the unstressed syllables which follow)” (Cruttenden, 1997; 20). Mortimer (1985; 77) calls the same thing a foot, and he counts an anacrusis as a foot. He supposes a “silent stress” at the beginning of an anacrusis, for example (an oblique stroke indicates the foot boundary and the mark + means a silent stress):

(a) /+There's a/woman./+In my/office./+And she/says she/wants to/  
see you./

If we use Cruttenden's framework, we can analyze (a) as following (capital letters indicate stressed syllables):

(b) /There's a WOman./In my OFfice./And she SAYS she/WANTS to/  
SEE you./

An anacrusis consists only of function words and, according to Cruttenden (1997; 21), syllables in an anacrusis generally tend to be produced “with greater speed than any unstressed syllables within following rhythm-groups.” The two syllable chunk of “there’s a” is likely to be reduced to one syllable and pronounced even without any vowel. This is an extraordinary characteristic of English especially for us Japanese speaking people. So we will pay special attention to it later through comparing stress time with mora time.

The theory of stress time or isochrony in English is closely related to the taxonomy of words. We classify them into content words and function words. The word stress in content words is fixed, while function words can be pronounced in two ways, that is, *strong forms* and *weak forms*. It is indispensable to teach weak forms together with stress time in order to improve Japanese students’ listening ability.

### **3. Stress time and mora time**

#### **3.1 Mora**

Japanese is sometimes called a syllable-timed language like French where every syllable is said to have nearly the same prominence. (See, Kreidler, 1989; 160) But there are several reasons against this view. One of them is that Japanese has no syllable, but mora.

A syllable consists of onset, nucleus and coda (onset and coda can be optional). Nucleus consisting of a vowel is indispensable to a syllable. On the contrary, Japanese has a phonological unit consisting only of a consonant.

According to Tsujimura (1996; 65), the mora in Japanese can be realized

in the following three cases:

- (1) (Consonant) Vowel (a consonant is optional)
- (2) the first part of a geminate (such as “sutoppu” meaning stop)
- (3) syllable final nasal /n/ (“moraic”/n/) (such as “hon” meaning a book)

This principle can be applied to any Japanese words which are divided according to the mora as follows: /su.to.p.pu/, /ho.n/.

It is the mora, not the syllable, that is a timing unit in Japanese. Each mora has the same length of time. That is to say, if a pair of three-mora words are pronounced, the durations of the two words are more or less the same. So we can say Japanese is a mora-timed language, while English is a stress-timed language. The difference between them is great.

First of all we cannot think of weak forms which are usually pronounced very quickly and ambiguously. Anacrusis are beyond our imagination. We are likely to treat each syllable as a mora and give an equal length of time to it. So we take much more time for pronouncing a stream of English speech than native speakers.

On top of that we have some trouble in pronouncing clusters of English consonants which are very rare in Japanese. There are only two cases where a kind of clusters appear in Japanese; one is the case of a geminate mentioned above, and the other one is the case of a consonant following the “moraic” /n/. For example, “hon” (book) + “ga” (the Case particle). Except these two cases, a consonant always appears before a vowel forming a (C)V type of mora.

So we pronounce English clusters as if each of the consonants made a

mora. For example, we pronounce “tasks” as if it were “tasukusu”.

Japanese has no diphthong. A seeming diphthong like “ai” (love) makes two morae and a seeming triphthong like “aoi” (blue) makes three morae. So when we pronounce the English name “Ian”, we are very likely to do so as three morae [I.a.n].

### 3.2 Rhythm and tone

Time is not the only problem we have when we pronounce a stress-timed language. Japanese has no rhythm as stress time in English (see, 2.3.2).

Surely, we have tone, but it is a part of intonation.

In the case of Japanese, each mora in a word is given a specific pitch. So the pitch pattern of the entire word is predictable given the location of the stress of a word. There are two rules of giving each mora its specific pitch. The first rule is that we give a high pitch to both of the stressed mora and the morae preceding it and give a low pitch to the morae after the stressed mora. The second one is called the initial lowering rule, that is, the pitch of the first mora of the word is low unless the stress is habitually given to that mora. (See, Tsujimura, 1996; 74–78)

Speaking a pitch-accent language like Japanese, if a word is pronounced with the wrong tone or intonation, then the word could mean an entirely different thing. For example (the mark /‘/ indicates stress):

[ni.n.‘ki] (popularity)——[‘ni.n.ki] (term of office)

[‘ki.ge.n] (period)——[ki.‘ge.n] (temper)

Thus, Japanese has tone or intonation, but no rhythm as stress time. The following two facts may make the non-rhythmical nature of Japanese clearer.

(1) In Japanese there are a large number of words that do not have stress, such as “ame” (candy), “aki” (vacancy), “hai” (ashes), “kan” (sense). In contrast, all content words in English have at least one stress to indicate prominence, and they sometimes have even a secondary stress like [‘edu’cation]; which is never found in Japanese.

(2) Disappearance of stress or prominence often happens in Japanese content words which have stresses outside the sentence context. For example, “naku” (chirp) has the low-high (LH) intonation as an independent word, but the second mora loses its prominence when the word is used in a sentence like “tori ga naku” (a bird chirps). In the same way, “aoi” (blue) with the LHL intonation becomes flat in a sentence like “yama ga aoi” (the mountains are blue).

Japanese is a overwhelmingly flat language in its intonation and has no rhythm like English. These characteristics make a great barrier to our pronunciation of a stress-timed language like English.

#### **4. Conclusion**

The suprasegmental features of English are pitch, duration and loudness. Exploiting these features, we produce the suprasegmental systems such as tone, intonation, rhythm and stress. In the case of English, tone is a part of intonation; rhythm is based on stresses and is called stress time.

If we teach the suprasegmental features of English systematically, we must focus our efforts on teaching intonation and rhythm or stress time. This is because they are so important in spoken English that some knowledge about them and productive practice of them are indispensable for improving students’ listening comprehension of a stream of speech in English.



Spoken English is especially difficult for Japanese students to understand. This is because the Japanese language has no rhythm as stress time; its intonation is rather flat; and it is a mora-timed language. If they pronounce English, they will take much more time than native speakers and their pronunciation will be flat and not rhythmical. These inborn habits of theirs are deeply related to their incomprehensibility of a stream of speech in English. Some knowledge about mora time and a systematic introduction to English intonation and rhythm, however, may help them improve their listening comprehension.

## 5. References

- Avery, Peter, and Ehrlich, Susan (1992) *Teaching American English Pronunciation*. Oxford: Oxford University Press.
- Brazil, David (1994) *Pronunciation for Advanced Learners of English*. Cambridge: Cambridge University Press.
- (1995) *Grammar of Speech*. Oxford: Oxford University Press.
- (1997) *The Communicative Value of Intonation in English*. Cambridge: Cambridge University Press.
- Cruttenden, Alan (1997) *Intonation* (Second Edition). Cambridge: Cambridge University Press.
- Dalton, Christiane, and Seidlhofer, Barbara (1994) *Pronunciation*. Oxford: Oxford University Press.
- Gimson, A. C. (1962 and 1989) *An Introduction to the Pronunciation of English* (Fourth Edition). London: Edward Arnold.
- Clark, John, and Yallop, Colin (1990 and 1995) *Phonetics and Phonology* (Second Edition). Oxford: Blackwell.
- Jones, Daniel (1909 and 1990) *The Pronunciation of English*. Cambridge: Cambridge University Press.
- Kreidler, Charles W. (1989) *The Pronunciation of English*. Oxford: Blackwell.
- (1997) *Describing Spoken English*. London: Routledge.
- Mortimer, Colin (1985) *Elements of Pronunciation*. Cambridge: Cambridge University Press.
- Nagano, Tadashi (1995) *Nihongo no Onseihyogen* (Speech Communication in

- Japanese). Tokyo: Tamagawa University Press.
- Nasr, Raja (1997) *Applied English Phonology*. Lanham: University Press of America.
- O'Connor, J. D. (1973 and 1991) *Phonetics*. London: Penguin Books.
- Pierrehumbert, Janet B., and Beckman, Mary E. (1988) *Japanese Tone Structure*. Cambridge, Massachusetts: The MIT Press.
- Pennington, Martha (1996) *Phonology in English Language Teaching*. London: Longman.
- Roach, Peter (1991) *English Phonetics and Phonology* (Second Edition). Cambridge: Cambridge University Press.
- (1992) *Introducing Phonetics*. London: Penguin English.
- Saitou, Yumiko (1990) *Nihongo Onseihyogenhou* (Speech and Recitation in Japanese). Tokyo: Oufu.
- Shimizu, Katumasa (1995) *Eigo Onseigaku* (English Phonetics). Tokyo: Keiso Shobo.
- Taylor, Linda (1993) *Pronunciation in Action*. New York: Prentice Hall.
- Tench, Paul (1996) *The Intonation Systems of English*. London: Cassell.
- Tsujimura, Natsuko (1996) *An Introduction to Japanese Linguistics*. Oxford: Blackwell.

(おおくま・さかえ 法学部教授)