



Suprasegmentals

In tone languages, it is usually relatively easy to see the function of the different tones. However, in languages which do not use tone in this way, it is harder to explain what we are doing when we make use of changes in pitch, loudness, and other **suprasegmental features** of speech. There are several such features: what they all have in common is that we usually see them as properties of pieces of speech which will be at least as long as one syllable, and may extend for many words. So if, for example, I say 'no' loudly, it is most likely that both the /n/ sound and the /əʊ/ will be loud. If I say 'hurry' quickly, then all the phonemes of that word will be said quickly. The most important suprasegmental features of speech are pitch, loudness, **tempo** (that is, speed), and **voice quality**, but these are by no means the only ones. The study of these features is often referred to as the study of **prosody**. Two such features form the basis for specially important functions, **stress** and **intonation**.

Stress and accent

In any language you listen to, you will notice that some syllables sound stronger and more noticeable than others. In English, for example, the middle syllable of the word 'tomato' is clearly stronger than the first and last syllables. We say that the middle syllable is *stressed*. In some languages the position of the strongest stress may affect the meaning of a word. The following Spanish words are shown with the stressed syllables underlined, and their meanings are given:

término 'terminus'
termino 'I terminate'
terminó 'I terminated'

English has some pairs of semantically related words whose grammatical category is reflected in their stress pattern, such as:

import (noun)
import (verb)

There are other pairs of words where the difference in stress signals functions in other ways: in the pair 'subject'/'subject', the two words differ in grammatical category (noun/verb) but seem in some contexts to be unrelated semantically; the pair 'recall'/'recall', however, may have the same grammatical category (both may be nouns) and be semantically related. In some other languages, it is possible to hear the difference between stressed and unstressed syllables, but the stress usually falls in the same position in a word of more than one syllable. In French, it is usual for stress to fall on the final syllable of the word, while in Polish it is usually on the penultimate syllable (the syllable before last), with a few exceptions such as 'uniwersytet' ('university'). In languages such as these, we cannot say that stress is able to determine the meaning of a word. We may guess that stress performs a different function: it helps us to divide the continuous flow of speech into separate words. For example, if I am listening to someone speaking French, I know that when I hear a stressed syllable, that is the last syllable of a word. One of the great unsolved mysteries of speech perception is how we manage to divide continuous speech up into separate words in languages like English (in which stress gives us relatively little help).

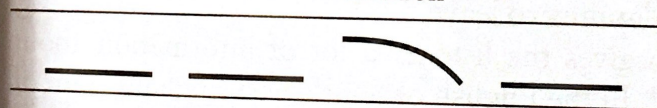
A number of factors cause a syllable to be made prominent so that it is heard as stressed. In English, stressed syllables are usually louder and longer than unstressed syllables, and have distinctive pitch (as in the example of the word 'important' given in the previous chapter). We can detect different levels of stress in words of several syllables. Try saying to yourself the four-syllable English word 'understanding': the strongest stress should be heard on the third syllable, but the second and fourth syllables are much weaker than the first syllable. Usually, only the third syllable has a noticeably distinct pitch.

Distinctive pitch, such as we find on the third syllable of 'understanding', is given special importance in the study of stress in English, and the term **accent**, which was introduced in the previous chapter, is used to refer to it. Consider now how you would say the phrase 'understanding English': you will probably find that there is no longer a noticeable pitch movement on the third syllable of 'understanding', but there is one on the first syllable of 'English'. The 'stand' syllable is still quite prominent, but it isn't accented. The word *stressed* can be retained for syllables (such as 'un' and 'stand' in this example) which are made prominent by other features such as length or loudness. This property is known as **stress**. The same distinction can be made in most other languages in which stress and accent play linguistically important roles.

Intonation

Intonation has always been a difficult thing to define. According to traditional descriptions, intonation is 'the melody of speech', and is to be analysed in terms of variations in pitch. We have seen, in the chapter on tone, how changes in pitch can change meaning, but in the case of intonation the way meaning is changed is not so clear. If we look at a typical example, we would expect a falling pitch pattern on a statement like this:

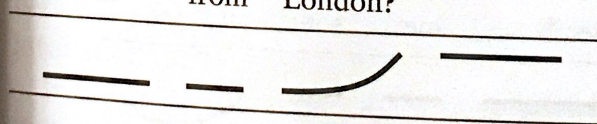
You're from London



The diagram shows a horizontal line representing a constant pitch level. Below this line, there are three short horizontal bars representing the duration of the words 'You're', 'from', and 'London'. A curved line starts at the beginning of the 'London' bar and curves downwards, ending at a lower pitch level than it started at, indicating a falling intonation pattern.

but a rising pitch pattern if the same words are used as a question:

You're from London?



The diagram shows a horizontal line representing a constant pitch level. Below this line, there are three short horizontal bars representing the duration of the words 'You're', 'from', and 'London?'. A curved line starts at the beginning of the 'London?' bar and curves upwards, ending at a higher pitch level than it started at, indicating a rising intonation pattern.

Intonation can, then, indicate different types of utterance, such as statements and questions. Other examples of meaning being changed by differences in intonation are often quoted: the difference between

She won't go out with anyone

and

She won't go out with anyone

is that the first one (with a falling pitch movement on 'any') says that she will go out with nobody, while the second (with a falling-rising pitch movement) says that she is careful about who she goes out with. In the case of

I have plans to leave

I am saying that I have some diagrams or drawings to leave, while

I have plans to leave

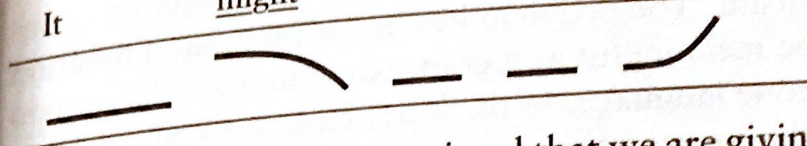
means that I am planning to leave.

Intonation also gives the listener a lot of information about what is being said. In the English of South-East England, a lot of use is made of a falling-rising intonation pattern: for example, it is often used for polite requests:

Can you lend me some money?

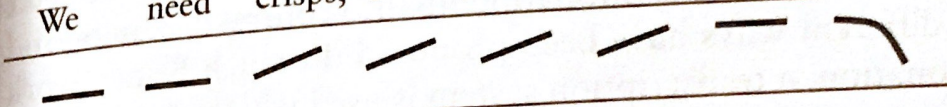
and for expressing reservation:

It might be good enough



We can use intonation to signal that we are giving a list:

We need crisps, drinks, fruit, sweets, and cheese



Intonation is said to indicate the attitudes and emotions of the speaker, so that a sentence like 'I think it's time to go now' can be said in a happy way, a sad way, an angry way, and so on. While this is certainly true, there is clearly more than just pitch variation involved in conveying such things, so the definition of intonation becomes considerably more complicated. It is clear that when we are expressing emotions, we also use different voice qualities, different speaking rates, facial expressions, gestures, and so on. What can we conclude about the use of intonation in a language like English? Although examples like those above can be produced which suggest functions related to the grammar of an utterance and the psychological state of the speaker, these seem to show only part of the picture. Perhaps the best way to look at the subject is to see intonation as an essential component of the **discourse** structure of speech. We speak in order to communicate, and we need to interact with our listeners to do this. We must indicate what type of information we are presenting and how it is structured, and at the same time we must keep our listeners' attention and their participation in the exchange of information. Communicative interaction would be much more difficult without intonation: think how many misunderstandings between people arise in the exchange of e-mail messages, where intonation cannot play a role.

It is difficult to work out a satisfactory way of transcribing intonation. In the examples given above, a 'wiggly-line' representation of the pitch movement is given, but although this helps to

explain the examples, it does not indicate which bits of intonation are significant. The problem lies in the fact that intonation is claimed to be meaningful and contrastive in a similar way to the tones of a tone language. In the latter case, we know that we can represent the tones with a set of marks which correspond to the contrastive tones of the language. But it is much more difficult to decide how to divide intonation up into contrastive units, and many different ways have been proposed. In much British work on intonation, a transcription system is used which places marks in the text to indicate the important features:

We need / crisps / drinks / fruit / sweets and \ cheese

Rhythm

There are many parallels between speech and music, and one thing that is always found in music is **rhythm**. In music, the rhythm is usually produced by making certain notes in a sequence stand out from others by being louder or longer or higher. We should not make the mistake of thinking that musical rhythm is just an unvarying repetition of beats at equal intervals. This may be true of commercial pop music (as can be heard coming out of someone's headphones, or through the wall from the room next door), but throughout the world in traditional folk music and other serious musical forms we can find some amazingly complex rhythms which are still immediately recognizable as regular. In speech, we find that syllables take the place of musical notes or beats, and in many languages the stressed syllables determine the rhythm. If you were asked to clap your hands in time with the sentence

'This is the 'frist 'time I've 'ever 'eaten a 'chocolate 'caterpillar

you would be most likely to clap at the points marked with the stress mark '. It is often claimed that English speakers try to keep an equal time between the stressed syllables, so the time between claps of your hands would be quite regular. Rhythm of this type is called **stress-timed**, and it is claimed that the unstressed syllables between the stressed syllables are squeezed into the time available, with the result that they may become very short. In fact, this

is only found in a style of speech (slow, emphatic) where the rhythm is strong, and in ordinary conversational speech it is much harder to make a convincing case for this **isochronous** rhythm (where the time intervals between stressed syllables are equal); as with music, we should not expect rhythm to be simple. Other languages have different rhythms (as you can easily hear by listening to them). To the ears of English speakers, Italian and Swedish have a very different rhythm from English. Spanish, French, and Chinese sound **syllable-timed** to English-speaking listeners—it sounds as though all the syllables are of equal length, and the dominant role of stressed syllables in making up the rhythm is much less noticeable. But these judgements are very subjective, and finding scientific evidence about what makes us hear languages as rhythmically different is proving to be very difficult. What does seem to be clear is that rhythm is useful to us in communicating: it helps us to find our way through the confusing stream of continuous speech, enabling us to divide speech into words or other units, to signal changes between topic or speaker, and to spot which items in the message are the most important.

Other suprasegmental features

It is possible to analyse the suprasegmental side of speech in great detail and to discover more features that vary as we speak. Some have been mentioned already: the speed or **tempo** of speaking is certainly something that we all can and do change, as is the loudness. **Voice quality** may vary from soft and gentle to harsh and unpleasant. Many of these features of speech are believed to be important in communication, but to have only a weak connection with the linguistic structure and phonology of the language since it is almost impossible to decide what is contrasted with what. They are therefore sometimes called **paralinguistic features**. These are very interesting in the study of how we behave when we speak to other people, particularly in the case of the expression of emotions. Even if you can't see the person who is speaking, you can detect emotions like anger, fear, happiness, or disgust in their voice. If we want to understand how this works, it certainly won't be sufficient to describe only the movements of the speaker's

pitch, as traditional intonation textbooks have done. We need a much more complex framework of features and descriptive labels, covering all aspects of the voice. At present we simply do not know in any detail what is happening when a particular emotion is conveyed, despite all the research being carried out.