

CHAPTER TWO

What is unique to listening

Introduction

In Chapter 1 I reviewed work from a number of academic disciplines, mainly linguistics and psychology, to arrive at a broad overview of the listening comprehension process. I concluded that listening comprehension is an active process of constructing meaning, and that this is done by applying knowledge to the incoming sound. I further concluded that comprehension is affected by a wide range of variables, and that potentially any characteristic of the speaker, the situation or the listener can affect the comprehension of the message.

Many of the important characteristics of listening comprehension are actually characteristics of all forms of language comprehension, and while listening ability is unique in some respects, it also shares many characteristics with reading. This should be clear from the discussion in Chapter 1, but there is also research evidence for this conclusion. For example, Buck (1992a) and Bae and Bachman (1998) have established by statistical means that there is considerable overlap between listening and reading ability, while both also have unique aspects. Freedle and Kostin have examined the sub-skills involved in processing a variety of multiple-choice test items, and comparison of the listening sub-skills (Freedle and Kostin, 1999) with the reading sub-skills (Freedle and Kostin, 1994) supports a similar conclusion.

This has practical implications for test development. It will become

clear in later chapters that testing listening is technically more complicated, more time consuming and far less convenient than testing reading: providing good quality recorded sound, for example, is just not as easy as handing out pieces of writing. If the purpose is to assess general language comprehension, it will usually be far easier to do that through reading. In many cases, it is not worth going to all the trouble of testing listening unless we are particularly interested in the knowledge, skills and abilities that are unique to listening.

However, listening is an important skill, and due to the practical complexities of providing spoken texts, it is neglected in many language-learning situations. One important reason to test listening, even when it might overlap quite considerably with reading, is to encourage teachers to teach it. In such cases, it is probably still better to emphasise the knowledge, skills and abilities which are unique to real-world listening, in order to encourage teachers to teach learners to comprehend realistic spoken language in a realistic manner.

The chapter will be organised under three main headings:

- i the important characteristics of spoken texts;
- ii the differences between first- and second-language listening;
- iii some common taxonomies of skills used in listening.

The important characteristics of spoken texts

In this section we will look more closely at those aspects of comprehension that are unique to listening. It is my belief that these are the characteristics we need to pay particular attention to if we want to create good listening tests. We will look at phonology, accents, prosodic features, speech rate, hesitations and discourse structure.

Phonological modification

The sounds of a language must be learned in order to understand speech. This is obvious; it is not the sounds themselves, however, that cause the most comprehension problems, but the way they vary in normal speech. In rapid speech, adjacent sounds influence each other. Such modifications take place according to a set of very complex rules, and these rules vary from one language to another. A

good description of phonological modification in English, and its relevance for second-language education, is provided by Roach (2001).

The degree of phonological modification varies depending on the situation. For example, in formal situations speakers will tend to have less modification than in informal situations; similarly, in cases where the information is important, they will tend to pronounce the words with more care than they would with casual, throw-away information.

In English the most important phonological changes are:

Assimilation, when sounds influence the pronunciation of adjacent sounds. For example, *'won't you'* is generally pronounced something like *'wonchoo'*.

Elision, when sounds are dropped in rapid speech. For example, *'next day'* is usually pronounced something like *'nexday'*.

Intrusion, when a new sound is introduced between other sounds. For example, in standard British English the sound /r/ at the end of the word *'far'* is not normally pronounced, but if the word is immediately followed by a vowel, as in *'far away'*, then it is inserted between the two words.

Also in English, the little words with a grammatical function usually have two pronunciations, a **strong form**, which is used in isolation or when the word is receiving stress, and a **weak form**, which is used in connected speech when the word has no sentence stress.

The modifications to pronunciation that take place during fast speech, especially informal speech, are quite extensive. Hardly a word is left unaffected. Other languages also have phonological modification, but according to a different set of rules.

Phonological modification and language comprehension

The phonological system – the complex set of rules that determine the pronunciation of connected speech – must of course be learned. Any lack of such knowledge is likely to be reflected in reduced comprehension. Many students who have studied a foreign language are shocked when they go abroad and hear native speakers for the first time, because the pronunciation encountered in the classroom often differs considerably from that used by real speakers. In a research

study, Henricksen (1984) found that most native speakers had little difficulty understanding words in their modified form, whereas the presence of phonological modification significantly reduces comprehension for second-language listeners. Even higher-level L2 listeners can fail to recognise language they actually know very well. Listeners need to know how the sound system works, in order to be able to process natural speech in real time.

Accent

Another occasion when listeners will encounter words pronounced in a non-standard manner is when listening to speakers with unfamiliar accents. It is very normal for different groups of language users to pronounce language in characteristic ways, and everyone has an accent. The most common accents are related to geography. In the case of English, Australians pronounce English differently from Americans, and this is different again from British English. Even within countries there can be considerable variation: the pronunciation in the north of England is very different from the south. Scotland and Wales have their own characteristic accents, as do Texas, Minnesota and New England in the United States. This is the case with most languages. Japanese is pronounced differently in the north than in Tokyo, and is different again in Osaka. Learners of Japanese as a second language will learn the standard Tokyo pronunciation, and will probably be unable to understand speech from rural areas in the provinces. Even Japanese native speakers from Tokyo or Osaka may not understand country folk from the far north or south. A similar thing happens with most languages that are dispersed across a geographical area, including French, Spanish, Arabic, Chinese and so on. In addition, accents not only vary according to geographical region, but also according to social groups. In England, for example, working-class English is different from middle-class English, which is different again from upper-class English.

Accent and language comprehension

Native speakers are generally used to hearing a wide variety of accents, but even then, when they hear a new accent for the first time,

they may not understand it very well, and it can take a little while for them to get used to it. L2 listeners are usually much less familiar with the range of common accents, and they sometimes have considerable problems when they hear a new accent for the first time (Kennedy, 1978). Eisenstein and Berkowitz (1981) showed that standard English is more intelligible to ESL learners than either the local working-class accent of New York, or foreign-accented English, and Anderson-Hsieh and Koehler (1988) found that the stronger the accent, the lower the listeners' comprehension. It seems reasonable to suppose that it generally takes L2 learners much longer to adjust to a new accent than native speakers.

Accent is potentially a very important variable in listening comprehension. When listeners hear an unfamiliar accent – perhaps hearing an Australian for the first time after studying with American teachers – this can cause problems and may disrupt the whole comprehension process. An unfamiliar accent can make comprehension almost impossible for the listener.

Prosodic features

Understanding the sound system of a language involves far more than just knowing the pronunciation of individual sounds and how they change in rapid speech. The prosodic features, stress and intonation, also carry a great deal of communicative information. How this works varies from language to language (see Cruttenden, 1997).

Stress

Stress is very important. In English, stressed syllables are not only louder, they are usually more clearly enunciated, longer, and are often preceded or followed by a short pause. There are two types of stress which are important in English, word stress and sentence stress.

Word stress is the relative emphasis of the various syllables within a word, and it forms a very important part of the phonological shape of that word. In Japanese all syllables generally have the same stress, but in English they usually get different amounts of stress. Even if an English word is pronounced with the correct sound sequence, it will

often be misunderstood if the relative stress of the syllables is incorrect. This is especially important with longer words.

Sentence stress is the relative emphasis of the words within an utterance. In English the most important words in an utterance are given more stress than the surrounding words. Words are stressed to indicate the point the speaker is making. For example, if a speaker says:

my SISTER returned yesterday

the stress is on the word 'sister', indicating that the topic of the utterance is the person, rather than what she did. However, if the stress changes to

my sister RETURNED yesterday

the topic is what she did, rather than who did it.

Another characteristic of stress, is that in some languages it determines the pace of the speech. This is the case in English, which is a **stress-timed** language. What this means is that the time between stressed syllables remains reasonably constant in any utterance. In the following, the words 'fax' and 'now' are stressed, and the time between them is the same in each utterance, and each utterance would take the same amount of time (Van Lier, 1995).

FAX NOW

FAX it NOW

FAX it right NOW

The result of this is that the words between the stressed syllables are pronounced very quickly, with no stress, and if there are more words they are pronounced even quicker to fit in the short time. To fit them in so quickly they are subject to considerable phonological modification, and sometimes may almost disappear.

This is a characteristic of English. In other languages the timing is determined in different ways. Japanese, for example, is a syllable-timed language: each syllable occupies the same amount of time regardless of how many phonemes there are in the syllable.

Intonation

Intonation is the variation in pitch that takes place within an utterance. Generally, in English, statements end with a falling intonation

and questions with a rising intonation. The intonation can make a considerable difference to the meaning.

Intonation has a number of very important functions. According to Crystal (1995) the most important ones are:

- **Emotional:** the intonation is used to express the speaker's attitudinal meaning, such as enthusiasm, doubt, or distaste.
- **Grammatical:** the intonation is used to mark the grammatical structure of the utterance, rather like the punctuation in written language.
- **Informational:** the intonation indicates the salient parts of an utterance, rather like sentence stress, so a higher pitch marks the important information.
- **Textual:** the intonation is used to help larger chunks of discourse to contrast or cohere, rather like paragraphs in written language.
- **Psychological:** the intonation is used to chunk information into units which are easier to deal with. For example, lists of words, or telephone and credit card numbers are grouped into units to make them easier to memorise.
- **Indexical:** the intonation is used by certain groups of people as a sort of identifier. For example preachers and newscasters often use a recognisable intonation pattern.

Intonation patterns do tend to vary from one language group to another, and they also change over time. A new intonation pattern is spreading across the English speaking world at present. This is a rising intonation at the end of statements. Women use it more than men, young people more than older people, and working-class more than professional-class people. Crystal (1995) suggests that it seems to have started in Australia or New Zealand, but I have heard it frequently among young people, especially women, in the United States. Perhaps in a few years it will become standard.

Prosody and comprehension

Stress and intonation are very important in word recognition. Dirven and Oakeshott-Taylor (1984) claim that they are more important to word recognition than the actual sounds, and Gimson (1980) suggests that 'a word pronounced with the correct sound sequence may well

be misunderstood if the relative prominence of the syllables is incorrect.’ Among their collection of **slips of the ear**, when people thought they heard something that was different to what was actually said, Bond and Garnes (1980) found that listeners usually understood the prosodic features of words, even when they did not catch the actual sounds.

Word stress, sentence stress and intonation are all important in comprehension. Lynch (1998) suggests that prosodic features have a direct impact on how listeners chunk and interpret discourse segments. Furthermore, these prosodic features can carry considerable meaning that supplements, or in some cases contradicts, the literal meaning of the words. Much of the communicative effect of utterances is expressed by the stress and intonation, and listeners need to be able to understand that to construct a reasonable interpretation.

Speech rate

All second-language listeners have probably had the experience of listening to something and not quite understanding it because it seemed too fast, of feeling that they could have understood if only it had been a little slower. Listener perceptions that speech is too fast are often due to a lack of processing automaticity, and as listeners get better, and as they learn to process the language more automatically, speech seems to become slower. So just because speech appears fast, it does not mean it is. However, the actual speech rate does affect comprehension, and there is a whole body of research that looks at the relationship between speech rate and comprehension. Results generally support the common-sense belief that the faster the speech, the more difficult it is to comprehend.

Typical speech rates

It is useful to start by getting an idea of typical speech rates. Tauroza and Allison (1990) looked at average speech rates for British speakers: radio monologues, conversations and interviews aimed at native speakers, as well as lectures intended predominantly for a non-native speaking audience. Table 2.1 summarises their findings.

The most common measure of speech rate is words per minute (wpm), which is obviously problematic when we consider that some

Table 2.1. *Average Speech Rates for British English (Tauroza and Allison, 1990)*

Text Type	words/ minute	syllables/ minute	syllables/ second	syllables/ word
Radio Monologues	160	250	4.17	1.6
Conversations	210	260	4.33	1.3
Interviews	190	250	4.17	1.3
Lectures to NNS	140	190	3.17	1.4

words are much longer than others, and further that there are systematic differences in word length between text types. Table 2.1 shows how word length varies, with the conversations and the interviews having a shorter word length, 1.3 syllables per word, whereas the monologues have 1.6 syllables per word. It is quite likely that serious academic lectures would have had even longer average word length. Syllables are a much better unit of measurement whenever precision is necessary (Sticht, 1971; Tauroza and Allison, 1990). However, counting syllables is complex and time-consuming, whereas counting words is relatively easy (especially with a word processor), so words per minute continues to be widely used.

Looking at the speech rates, we can see that the average is about 170 wpm, or about 4 syllables per second (sps). Interactive speech – the conversations and the interviews – is a little faster, and monologues are a little slower. The lectures to the non-native speakers are much slower, at about 140 wpm or 3.17 sps, which suggests that the lecturers are probably slowing down their speech to aid comprehension. It is important to remember that these are averages, and that there will be considerable variation between different speakers: the speaker in Example 2.2, later in the chapter, normally speaks at about 230 wpm, which is very fast.

Relationship between speech rate and comprehension

What is important for us is the relationship between speech rate and comprehension. Looking first at the situation with native speakers, in an early study Foulke (1968) modified the speed of recordings from a talking book and found that comprehension was relatively unaffected by increases up to 250 wpm, but above this point it began to drop off

rapidly. Foulke and Sticht (1969) and Sticht (1971) reviewed a number of other studies and concluded that there is indeed a gradual decline in comprehension, until a threshold level is reached at about 275 wpm, after which comprehension drops off far more steeply. Carver (1973) found similar results and interpreted them as follows:

The data suggest that when an individual is presented with speech at increasingly high rates, there will be little decrease in his understanding of the thoughts until the minimum amount of time required to process each word, each group of words, or each thought is exceeded. Then, the decrease in thoughts understood will drop abruptly because little or nothing is understood or comprehended below the threshold. (Carver, 1973:124)

Anderson-Hsieh and Koehler (1988) looked at speech rate and foreign accent. Predictably they found that native English speakers had lower comprehension at faster speech rates, but more interestingly, they found that the decrease in comprehension was most dramatic when the speaker had a more pronounced accent. This suggests that speech rate is more critical to comprehension when listening to speakers with less standard accents.

There is less research on the effects of speech rate on second-language comprehension. Stanley noted that the thing that gave listeners most difficulty when listening to everyday speech was 'quite simply, those areas which were spoken most rapidly' (1978:287). He adds that when speech was faster, language learners 'constantly failed to perceive individual phonemes and hence words with which they were already familiar' (1978:289). Griffiths (1992) looked at the effects of three different speech rates (127, 188, 250 wpm) on the comprehension of second-language speakers and concluded that comprehension was significantly better at the slowest speech rate and worse at the higher rates.

To summarise, the research has shown speech rate to be clearly an important variable in listening comprehension. Comprehension declines as the speaker talks faster, and the weight of the evidence suggests that the decline in comprehension is rather slow until a threshold level is reached, at which time an increased speech rate leads to a much more rapid decline in comprehension. As for the particular speech rate at which this rapid decline takes place, the weight of the evidence suggests that this varies from one listener to another and is affected by a variety of factors. There is strong evidence to suggest that the language ability of the listener is important

– non-native speakers seem to have a lower threshold – and that the accent of the speaker is also important. Based on this, we can probably assume that other text variables, such as vocabulary, syntax or topic, would also interact with speech rate.

Hesitation and language comprehension

One important characteristic of spoken texts are hesitations. There are four main types of hesitations: (i) unfilled pauses, which are just periods of silence; (ii) filled pauses, where the speaker uses fillers such as *'uh'*, *'um'*, *'ah'*, *'mm'* or *'well'*, *'anyway'*, *'let me see'*; (iii) repetitions, where the speaker repeats the same word or part of a word; and (iv) false starts, where the speaker stops and then replaces the previous word or phrase with another choice.

There is evidence to suggest that hesitation phenomena can present a major comprehension difficulty to non-native speakers who are listening to spontaneous speech. Voss (1979) asked 22 non-native speakers of English to listen to a passage of spontaneous speech, about 210 words long. The passage included a variety of hesitation phenomena: repeats, false starts, filled pauses and unfilled pauses. Results indicated that nearly one-third of all perception errors were connected with hesitation phenomena. These errors were due to listeners either misinterpreting hesitation phenomena as words, or parts of words, or to misinterpreting parts of words as hesitations.

However, there is also evidence to suggest that certain types of hesitation phenomena can aid comprehension. When one-second pauses were introduced into a text at the juncture between clauses, comprehension improved, while pauses inserted randomly into the text resulted in lower levels of comprehension (Friedman and Johnson, 1971; Johnson and Friedman, 1971). Blau (1990) inserted pauses into a variety of passages and found that comprehension of the text with the pauses was significantly better. However, Blau's results could have been because the listening passages with the pauses were slower and hence easier to understand. In a follow-up study, Blau (1991) used a new version of the recording in which pauses were filled with hesitation markers. Results indicated that comprehension scores for both hesitation passages, that is with unfilled pauses and filled pauses, were significantly higher than for the normal passages, but did not differ significantly from each other.

To summarise, the research on hesitation phenomena is quite clear in indicating that these occur as a regular and normal part of spoken English. Hence the ability to comprehend spoken English must include the facility to deal with these hesitation phenomena. As for how they affect comprehension, there are indications that in some cases they can aid comprehension, and in others they may cause problems. But the research suggests that pauses inserted at meaningful syntactic boundaries can aid comprehension, whereas random pauses do not. Hesitations that slow down the speech rate do aid comprehension for L2 listeners, and this applies to both filled and unfilled pauses, as long as the listener recognises the fillers as fillers. However, if the listener does not recognise these as fillers, then comprehension can be adversely affected.

Discourse structure

In Chapter One I reviewed work on important aspects of discourse comprehension, most of it conducted on first-language reading. There is also a body of work that looks at the effect of discourse on second-language listening comprehension, the work on academic lectures. This is important for our purpose because a common reason for testing listening comprehension is to determine whether students are ready for academic study.

I will also illustrate aspects of spoken discourse structure by presenting transcripts of actual speech samples. However, there are some problems connected with this. When speech is transcribed, many of its important characteristics are lost. Most importantly it loses stress, intonation and timing, and as we have seen these convey a great deal of important information. However, writing speech down does have the important advantage of making it clear just how different it is from writing. Some of the following examples may be difficult to read, but they are normal speech, and they made perfectly good sense when they were spoken.

Lecture comprehension

Many researchers believe that discourse variables are very important in lecture comprehension. For example, Olsen and Huckin (1990)

point out that ESL students can understand all the words of a lecture and still fail to understand the main points, and Dunkel and Davis (1994) claim that lecture comprehension depends less on the meaning of the individual sentences, and more on their inter-relatedness and the structure of the whole text. Devices that explicitly signal the micro- and macro-structure of the text are assumed to be very important, and it is claimed that effective speakers make use of appropriate structuring and organising cues to help listeners understand the meaning of the lecture (Chaudron and Richards, 1986). Research suggests that both L1 listeners (Kintsch and Yarborough, 1982; Meyer and Freedle, 1984) and L2 listeners (Hron *et al.*, 1985; Chaudron and Richards, 1986) can benefit from explicit discourse markers within the text.

Lectures can be very challenging for second-language listeners, and some research has tried to discover how to make them easier: redundancy (using repetition and paraphrase) and the use of clear rhetorical markers are assumed to aid comprehension. Chiang and Dunkel (1992) found that added redundancy benefited higher ability students more than lower ability students, and Dunkel and Davis (1994) found that rhetorical signal cues did not aid comprehension.

Tauroza and Allison (1994) found that students have difficulty following more complex discourse structures. However, it is possible that lectures with the more complex structure also imposed a greater cognitive load on the listeners, and Brown (1995b) has argued that cognitive load is one of the most important determinants of difficulty in listening.

Speech sample one

The following is a transcript of a typical lecture, taken from Rost (1994). It is the beginning of an introductory lecture on social psychology.

Example 2.1

- 1 hello everybody. today i would like to continue our discussions of
- 2 social psychology. by talking about some of the major influences on
- 3 our attitudes and our behaviours. i think the place to start is by in-
- 4 troducing the idea or the concept of a norm. N-O-R-M. and I am
- 5 sure you are all familiar with this concept. a norm is an accepted or

6 expected standard of behaviour. or standard of thinking in a given
 7 group of people. and we find that all groups of people have norms.
 8 for an array of things. for basic things like the food they eat, the
 9 clothes they wear, to much more complex types of behaviour and
 10 decisions.

Although lectures are more formal than many other speech events, and although this is a very short extract, the discourse patterns are clearly different from written language. The speaker strings together direct statements, in a rather loose way, by just adding more information, and following on from one idea to the next. This is the most basic characteristic of discourse structure in spoken language.

Speech sample two

Academic lectures form a separate domain of language use, and are somewhat different from more informal speech. Example 2.2 is more typical of daily interaction. Here, a young American man is describing the time he arrived to take part in a tennis tournament and discovered that there was only one other person besides himself entered for the tournament.

Example 2.2

1 and i got there, and it turned out i was only the second person
 2 signed up for the entire tournament. and so there_ the people who
 3 ran it were like 'do you even want to bother playing the tournament?'
 4 and i was like_ 'let's it a great opportunity because if i win then i've
 5 won a tournament only with one match. so i can beat one guy, and
 6 even if i don't, like who cares.' so_ and_ cuz you get points for when
 7 you finish. so a runner-up gets the points anyway. so i was kind of
 8 hoping that the other guy i was gonna playing against was gonna
 9 dropout, and just give me the trophy. but he didn't. so anyway we
 10 went out and played, and we didn't even start_ we started like way
 11 after_ because i got there for a seven o'clock match. we didn't start
 12 playing until like ten at night.

There are no significant pauses in the original, but obviously the stress, intonation and timing make the structure of the passage quite clear. This passage illustrates some noteworthy characteristics. Firstly there are a number of false starts. In line 2, there is a repair after '*so there*', and later, on line 6 after '*so and*', he changes his mind and

offers an explanation of why he doesn't mind if he gets beaten. In line 10 too there is a repair after *'even start'*, and in line 11 one after *'like way after'*. He also appears to make a grammatical error in line 8, *'the other guy I was gonna playing against'*, although in rapid speech the two words *'playing'* and *'against'* run together and it is hard to be sure what was said.

Probably the most interesting aspect of this text, however, is the use of the word *'like'* to introduce a direct or indirect quotation, or even just a thought. He first uses this word when he explains how the organisers were asking the two players whether they actually wanted to play or not. In this case he uses it to introduce a direct quotation, *'the people who ran it were like do you even want to bother'*. He uses the same *'like'* to introduce his reaction, *'and i was like let's it a great opportunity'*. It seems likely that he is not necessarily quoting himself directly, but probably explaining his thoughts on the reasons that motivated him to continue with the match. He then uses *'like'* again in line 6, when he says, *'like who cares'*. In this case, however, this is clearly not a quotation, but an expression of his own attitude to the situation.

The point here is not to go into details of the usage of this word *'like'*, but to note that this is not a standard use of the word. Some would consider it slang, or think it incorrect usage, but the point is that this is very common among young Americans. This is not street slang, nor is this a conversation between two young teenagers. The speaker is an undergraduate at Princeton University, an Ivy League university with very high admission standards and, although the situation was relaxed and relatively informal, he is speaking to an older male who was a relative stranger.

Speech sample three

The third example is of a young American woman, a 19-year-old student at a state university. She is talking to a relative stranger, an older male, and telling him about how she got into serious debt with her credit cards.

Example 2.3

- 1 um now here comes August, and i'm thinking okay i'm getting bills.
- 2 they're saying eight hundred dollars, and six hundred dollars and

3 I'm like 'wow oh my goodness.' so yaknow i haven't stopped yet,
 4 though, because i have them_ i have_ i keep one_ i stopped_ i put
 5 two away_ i acquired another one mid- summer, used about three
 6 hundred dollars on that one. and . . . uh i had one more that i was
 7 still using because it was my A-T and T card, it was also a phone
 8 card, so it was very helpful. um but anyway, um i finally started
 9 using_ i i_ continue to use to that A-T and T card. and i saw little
 10 sales that just yaknow i couldn't resist. they were perfect sales and_
 11 they were good sales, but i didn't have money. it was_ using_ i was
 12 using the credit cards, so i finally yaknow i called one day, i tried to
 13 use the A-T and T card and it said no.

Again it is important to stress that this is normal informal speech and quite intelligible. Again there are no significant pauses in the original, but the rhythm, stress and intonation make the structure very clear. There are numerous fillers (e.g. 'um', 'uh'), repetitions (e.g. line 9: 'i i'), repairs (line 4: 'i have i keep one' or line 11: 'it was using i was using'), as well as exclamations ('wow!') and non-standard forms ('like', 'yaknow'). What is most interesting however, is the discourse structure. This is really a stream of ideas, connected not so much by grammatical devices, but more by their juxtaposition in the text and the coherence of the ideas. In fact on reading this it appears very ungrammatical.

This is what much spoken language is really like. Reading a theoretical description as in Chapter 1 may be helpful, but it does not have the same impact as actually seeing normal speech written down in this way. These are not extreme examples of the spoken genre – readers can be assured that a group of close friends, chatting informally over a cup of coffee, would likely produce language with many more oral characteristics than the examples here.

Non-verbal signals

Before concluding this section on the characteristics of spoken texts, it is important to note that in many spoken interactions, the relevant linguistic information is conveyed not only by the sound. Kellerman (1990) has argued that looking at the speaker's mouth – the lips, jaw and tip of the tongue – provides information about what the speaker is saying, and listeners use that to help them understand. And not all the relevant information is conveyed by the language. As Abercrombie

suggests, 'we speak with our vocal organs, but we converse with our bodies' (1967:55). Visual information is important in communication, and of particular interest given the increased availability of video and multimedia.

Work on teaching second-language listening suggests that visual support can aid language learners, especially less proficient learners, and is particularly helpful with more difficult texts (Rubin, 1995). However, whether visual information generally increases comprehension is still open to doubt, although test-takers certainly seem to prefer it (Progosh, 1996). Lynch (1998) suggests that it may increase motivation and attention levels, and comprehension of gist, but that it probably has little effect on understanding detail, and there is no firm evidence that it increases long-term retention.

Bostrom (1997) notes that the non-verbal signals sometimes contradict the verbal information, and in such cases listeners 'accept the non-verbal as a more valid expression of the feelings of the interactant' (Bostrom, 1997:24). There is research to indicate that listeners vary in their ability to understand these non-verbal cues (Burgoon, 1994), but Bostrom claims that visual cues are decoded with much greater accuracy than many verbal messages.

Non-verbal communications can take a number of forms, and they may be deliberate, for the purpose of communication, or unintentional. Firstly, there are actions or movements that are obligatory in certain types of social situations. For example, greetings such as bowing and handshaking are executed in a certain way, and in some social groups there is often a correct way to hold oneself while speaking to an employer or high-ranking superior officer. Refusing to behave in that respectful manner will have a very strong communicative effect, however polite one's words.

Secondly, certain general body movements express the mood of the speaker: depression or happiness are often very apparent in the posture and movements of the speaker, and our assessment of the speaker's mood can have a considerable influence on how we interpret what they say.

Thirdly, there is kinesics, which we might define as message-related body movements (Antes, 1996; Kellerman, 1992). Gestures or facial expressions can substitute for a verbal message, as in the case of a shoulder shrug to indicate that the speaker doesn't know or doesn't care. Sometimes body movements might reinforce a message, as in the case of emphatic gestures stressing important points, or at other

times they might be intended to completely modify the interpretation of the spoken message, as in the case of gestures or facial expressions to indicate disbelief or sarcasm.

Although the effect of visual information may vary considerably from one listening situation to another, it certainly has the potential to influence or change the listener's interpretation of the speaker's words in a significant way.

The differences between first- and second-language listening

The view of listening comprehension described in Chapter One was developed mainly through first-language comprehension. In the last section we looked at what is unique about listening compared to other types of comprehension. In this section we will look at the difference between first- and second-language listening.

Although there is not much work that looks explicitly at the difference between first- and second-language processing (see Lynch, 1998), the existing work suggests that the processes are similar. For example, Fishman (1980) found that both first- and second-language test-takers made similar types of errors. Voss (1984) looked at slips of the ear, and also found that both natives and non-natives made similar types of errors.

I believe that the evidence to date gives no reason to suppose that second-language listening is in any fundamental way different from first-language listening. The same processes seem to apply. Dunkel (1991a), for example, makes a similar assumption in her review of native- and second-/foreign-language listening. Faerch and Kasper (1986) claim that first- and second-language listening differ only in that the second-language listener has a restricted knowledge of the language, and that they might be influenced by transfer from their first language.

There is, however, a difference in emphasis. When problems arise in native-language listening, they are often due to such factors as the listener being distracted, disinterested or responding to the content by thinking about something else. While these will obviously be of importance to all listeners, with second-language listening more problems arise due to insufficient knowledge of the linguistic system, or a lack of knowledge of the socio-cultural content of the message.

They may lack crucial content or textual schemata to help them understand the text (Long, 1989; Chiang and Dunkel, 1992), and because they generally come from a different background, second-language learners often lack important background knowledge that would help them compensate for their lack of linguistic skills (see Aitchison, 1994; Bremer *et al.*, 1996).

First-language knowledge

All normal human beings successfully acquire their first language. In our first language, we are all capable of understanding complex and subtle ideas. Indeed, first-language listeners become very adept at detecting affective cues about the speaker's state of mind or attitude towards the message (Bostrom, 1997). This is not always the case with second-language listeners.

First-language knowledge is largely implicit, by which I mean that it is outside conscious awareness. It is procedural knowledge rather than declarative knowledge. We do not consciously learn most of our first-language skills, we just acquire them automatically as we grow up and interact with the world around us. In our first language, we all know many complex rules and can use them correctly without being aware of them, and usually we could not explain these rules if we tried. The cognitive processes that are involved in first-language comprehension are almost entirely automatic – it is as if the language were transparent: we just do not notice it is there. When we hear something in our first language, we just automatically understand.

Second-language knowledge

In the case of a second language, we usually learn things differently, especially when we learn it as an adult. Knowledge of the language varies considerably: from those who can understand only a few isolated words in their second language, to those who appear to be indistinguishable from native speakers. Rarely do second-language learners develop the same high level of ability as in their first language. In most cases, second-language listeners have huge gaps in their knowledge of the language. Although some second-language knowledge is implicit, much of the knowledge is explicit, by which I

mean that it is to some extent consciously known and consciously applied.

For example, in a study which examined dictation errors, Angela Oakeshott-Taylor (1977) found evidence that proficient subjects were able to process the sound quickly and derive meaning from what they were hearing, whereas weaker subjects had to spend so much time on basic linguistic analysis that they had no time for interpretation of the meaning. She concludes, 'the process that takes place with the native speaker in a matter of milliseconds may occur over a longer period of time with the less proficient L2 speaker' (1977:148).

If we think of language as a window through which we look at what the speaker is saying, in the case of first-language listening, the glass is very clean and we see through it without even noticing it is there; but in the case of second-language listening, the glass is dirty: we can see clearly through some parts, other parts are smudged, and yet other parts are so dirty we cannot see through them at all. We are very aware of the glass because it gets in the way.

When second-language learners are listening, there will often be gaps in their understanding. In some cases, these gaps will be just a small part of what they hear, an occasional word or nuance; but in other cases the gaps will form a significant proportion of the message, and in more extreme cases, the listener may only understand a few isolated words or expressions. Of course, gaps occur in first-language listening too, but the gaps in second-language listening usually have a far more significant impact on comprehension.

Compensatory skills

As we have seen, listeners use their understanding of the communicative situation – the speakers or the topic under discussion – to help them understand what is being said. Therefore, when there is a gap in their linguistic knowledge, second-language listeners will naturally tend to compensate for that by using any other available information – including visual information, general background knowledge or their common sense. A second-language listener who knows the situation well and who knows what to expect, will often be able to understand a great deal even when there are considerable gaps in their knowledge of the language. These compensatory skills are a significant aspect of second-language listening.

However, this may not always be so easy for second-language listeners who come from different backgrounds: what is significant in one cultural situation may not be significant in another. Similarly, non-verbal information may have different meanings from one culture to another. For example, in certain situations in Japan, it is common to laugh as an expression of extreme embarrassment; Americans would never laugh in such situations and so they could easily misinterpret this as indicating a lack of concern. The second-language listener may be at a considerable disadvantage, because the background knowledge necessary to understand and interpret the language may not be available.

I believe that the difference between first- and second-language listening is not that the processes are fundamentally different in any way, but only that the knowledge necessary to understand is often grossly inadequate for the second-language listener. This may often be a double disadvantage, in that they lack both the knowledge of the language and also the background knowledge to compensate for that.

Listening sub-skills

Listening comprehension is a complex, multidimensional process, and a number of theorists have attempted to describe it in terms of taxonomies of sub-skills that underlie the process. However, it should be stressed at the outset that the empirical support for these taxonomies is usually lacking. While there is no doubt that many of the components are crucial in listening, there is no evidence to suggest that any of these taxonomies constitute a complete unified description of the listening process. Collectively they are useful because they tell us what scholars in the field have come to think is important in listening comprehension.

The two-stage view

One very common taxonomy is the division of listening into a two-stage process: a first stage, in which the basic linguistic information is extracted, and then a second stage in which that information is utilised for the communicative purpose. This division of listening com-

prehension is a theme that occurs again and again in the literature, and for this reason I believe it has considerable value.

Probably the clearest example is Carroll (1972), who describes listening comprehension as a two-stage process: first the apprehension of the linguistic information contained in the message, and second the application of that linguistic information to the wider communicative context. Rivers (1966) suggests that listening to a foreign language may be analysed into two levels of activity: 'the recognition level, involves the identification of words and phrases in their structural inter-relationships' and 'the level of selection, where the listener is drawing out from the communication those elements which seem to him [*sic*] to contain the gist of the message' (1966:199).

In a very influential text on speech processing, Clark and Clark (1977) make the distinction between the 'construction process', which is the way listeners construct an interpretation of a sentence from the speaker's words, and the 'utilisation process', which is concerned with how listeners utilise this interpretation. In the construction process the listener understands the underlying propositions which the speaker meant to express, and then in the utilisation process the listener uses these to understand what the speaker intended to communicate by that. Oakeshott-Taylor (1977) proposes a similar distinction between what he calls 'micro-comprehension' which is the perception of a short section of a text, and 'macro-comprehension' which is the understanding of a text in its totality (Oakeshott-Taylor, 1977:94).

Clark and Clark (1977) make the point that the distinction between the two processes may not be very clear. They stress that this does not infer a serial or sequential model of listening, and that in actual practice it may not be possible for a listener to infer the meaning of the propositions without simultaneously being aware of the speaker's purpose in using them. There will obviously be feedback between them, and the two aspects of listening comprehension will be capable of interacting and influencing each other.

Despite some differences, these scholars seem to have arrived at similar conceptualisations of listening comprehension, and the fact that they use different terminology suggests that they have arrived at this understanding more or less independently. This adds considerable credibility to the two-stage view of listening.

A cognitive skills approach

Quite a different approach underlies a taxonomy from Valette (1977:20), which is interesting because she intended this to form the basis for developing listening tests. Although she makes no reference to Bloom, the categories do seem heavily influenced by Bloom's (1956) taxonomy of educational objectives in the cognitive domain. Her taxonomy has five levels; it does not attempt to describe the processing of listening *per se*, but describes a series of increasingly complex cognitive skills that can be used to show increasing facility with listening comprehension.

- **Mechanical skills:** the listener performs by rote memory, rather than by understanding. For example, perceiving differences between two or more sounds and making distinctions between them.
- **Knowledge of the language:** the listener demonstrates this by showing knowledge of facts, rules etc. This can be done by answering simple questions about the language.
- **Transfer:** the listener uses this knowledge in new situations, for example taking information from one source and using it for a new purpose.
- **Communication:** the listener uses the language as a natural vehicle for communication, for example when a message containing new information is understood.
- **Criticism:** the student analyses or evaluates a piece of language in terms of its effectiveness, appropriacy, style, tone etc.

Communicative approaches

Given the interest in communicative language teaching, it is not surprising that some scholars have attempted to describe listening skills in communicative terms. An early attempt to do this was Aitken (1978). His taxonomy goes beyond basic linguistic processing and takes into account a variety of skills which are concerned with relating the basic linguistic processing to the wider communicative situation. He does not claim that these constitute a complete description of the skills used in listening, but he does claim that the skills needed to process and comprehend speech include at least the following:

- Understanding the vocabulary and being able to guess the meanings of unfamiliar or unclear words from their context.
- Understanding the syntactic patterns, the morphological forms characteristic of spoken language, and following the discourse patterns of spoken language.
- Understanding the flow of stressed and unstressed sounds, as well as intonation cues and other cues of oral punctuation.
- Identifying the speaker's purpose.
- Drawing correct conclusions and valid inferences about the social situation, the speaker's intent or the general context.
- Recognising the speaker's attitude to the listener and the subject of their discussion.
- Identifying the techniques and rhetorical devices the speaker used to convey the message.

A later, and more comprehensive taxonomy of communicative listening sub-skills is provided by Weir (1993). He does not call this a taxonomy as such, but a checklist of operations that listening tests should require. Like Aitken, he makes it clear he does not regard this as a complete list of listening sub-skills.

Direct meaning comprehension

- listening for gist
- listening for main idea(s) or important information; and distinguishing that from supporting detail, or examples
- listening for specifics, including recall of important details
- determining a speaker's attitude or intention towards a listener or a topic

Inferred meaning comprehension

- making inferences and deductions
- relating utterances to their social and situational contexts
- recognising the communicative function of utterances
- deducing meaning of unfamiliar lexical items from context

Contributory meaning comprehension

- understanding phonological features
- understanding grammatical notions such as comparison, cause, result, degree etc.
- understanding discourse markers
- understanding the main syntactic structure of clauses or idea units
- understanding cohesion, especially reference
- understanding lexical cohesion, especially lexical set membership and collocations
- understanding lexis

Listening and taking notes

- ability to extract salient points to summarise the text
- ability to select relevant key points

It is interesting that neither Aitken nor Weir attempt to provide exhaustive lists of communicative skills, but rather seem content to point out the important skills that they think should be included on listening tests.

More detailed taxonomies

However, other theorists have taken a different view and tried to present a far more complete taxonomy of sub-skills. One of these is Richards (1983). He asserts that 'listening purposes vary according to whether learners are involved in listening as a component of social action (e.g. conversational listening), listening for information, academic listening (e.g. lectures), listening for pleasure (e.g. radio, movies, television), or for some other reason' (1983:228). He suggests that different lists of 'micro-skills' (1983:228) would be required for each of these, and his suggestions for conversational listening are given in Table 2.2, and for academic listening in Table 2.3.

Table 2.2. *Conversational Listening (Richards, 1983)*

1. ability to retain chunks of language of different lengths for short periods
2. ability to discriminate between the distinctive sounds of the target language
3. ability to recognise the stress patterns of words
4. ability to recognise the rhythmic structure of English
5. ability to recognise the functions of stress and intonation to signal the information structure of utterances
6. ability to identify words in stressed and unstressed situations
7. ability to recognise reduced forms of words
8. ability to distinguish word boundaries
9. ability to recognise the typical word order patterns in the target language
10. ability to recognise vocabulary used in core conversational topics
11. ability to detect key words (i.e. those which identify topics and propositions)
12. ability to guess the meaning of words from the contexts in which they occur
13. ability to recognise grammatical word classes
14. ability to recognise major syntactic patterns and devices
15. ability to recognise cohesive devices in spoken discourse
16. ability to recognise elliptical forms of grammatical units and sentences
17. ability to detect sentence constituents
18. ability to distinguish between major and minor constituents
19. ability to detect meanings expressed in differing grammatical forms/ sentence types (i.e. that a particular meaning may be expressed in different ways)
20. ability to recognise the communicative functions of utterances, according to situations, participants, goals
21. ability to reconstruct or infer situations, goals, participants, procedures
22. ability to use real-world knowledge and experience to work out purpose, goals, settings, procedures
23. ability to predict outcomes from events described
24. ability to infer links and connections between events
25. ability to deduce causes and effects from events
26. ability to distinguish between literal and applied meanings
27. ability to identify and reconstruct topics and coherent structure from ongoing discourse involving two or more speakers
28. ability to recognise coherence in discourse, and to detect such relations as main idea, supporting idea, given information, new information, generalisation, exemplification
29. ability to process speech at different rates
30. ability to process speech containing pauses, errors, corrections
31. ability to make use of facial, paralinguistic and other clues to work out meaning
32. ability to adjust listening strategies to different kinds of listener purposes or goals
33. ability to signal comprehension or lack of comprehension, verbally and non-verbally

Table 2.3. *Academic Listening (Richards, 1983)*

1. ability to identify purpose and scope of lecture
2. ability to identify topic of lecture and follow topic development
3. ability to identify relationships among units within discourse (e.g. major ideas, generalisations, hypotheses, supporting ideas, examples)
4. ability to identify role of discourse markers in signalling structure of a lecture (e.g. conjunctions, adverbs, gambits, routines)
5. ability to infer relationships (e.g. cause, effect, conclusion)
6. ability to recognise key lexical items related to subject/topic
7. ability to deduce meanings of words from context
8. ability to recognise markers of cohesion
9. ability to recognise function of intonation to signal information structure (e.g. pitch, volume, pace, key)
10. ability to detect attitude of speaker toward subject matter
11. ability to follow different modes of lecturing: spoken, audio, audio-visual
12. ability to follow lecture despite differences in accent and speed
13. familiarity with different styles of lecturing: formal, conversational, read, unplanned
14. familiarity with different registers: written versus colloquial
15. ability to recognise relevant matter: jokes, digressions, meanderings
16. ability to recognise function of non-verbal cues as markers of emphasis and attitude
17. knowledge of classroom conventions (e.g. turn-taking, clarification requests)
18. ability to recognise instructional/learner tasks (e.g. warnings, suggestions, recommendations, advice, instructions)

One very influential taxonomy of language skills is Munby (1978), who provides a very detailed taxonomy of what he calls 'enabling-skills'. This includes both productive and receptive, spoken and written skills, and provides the most exhaustive list of sub-skills, a list too long to be quoted here.

Research-based taxonomies

The taxonomies described above are based primarily on theoretical speculation, but there are a small number of taxonomies based on research. Using a newly developed research methodology, Buck *et al.* (1997) looked at 30 listening items from Part Three of the TOEIC: short dialogues followed by one multiple-choice comprehension

question. They found 14 abilities which they claimed were most important:

- the ability to process faster input
- the ability to process lower-frequency vocabulary
- the ability to process text with higher vocabulary density
- the ability to process more complex structures
- the ability to process longer segments
- the ability to process text with a higher information density
- the ability to scan short segments to determine listening purpose
- the ability to synthesise scattered information
- the ability to use redundant information
- the ability to use word-matching strategies
- the ability to resist superficial word associations
- the ability to recall (arbitrary) names
- the ability to make text-based inferences
- the ability to use background knowledge to make inferences

In a similar study, Buck and Tatsuoka (1998) looked at a test with 35 short-answer comprehension questions, taken by a group of Japanese test-takers. They found 15 abilities that accounted for most of the test-taker performance:

- the ability to identify the task by determining what type of information to search for in order to complete the task
- the ability to scan relatively fast spoken text, automatically and in real time
- the ability to process a relatively large information load
- the ability to process a relatively medium information load
- the ability to process relatively dense information
- the ability to use previous items to help information location
- the ability to identify relevant information without any explicit marker to indicate it
- the ability to understand and utilise relatively heavy stress
- the ability to process relatively fast text automatically

- the ability to make text-based inferences
- the ability to incorporate background knowledge into text processing
- the ability to process L2 concepts with no literal equivalent in the L1
- the ability to recognise and use redundant information
- the ability to process information scattered throughout a text
- the ability to construct a response relatively quickly and efficiently

All the taxonomies in this section should be treated with caution. Those based on theory are only lists of what scholars think are likely to be important, whereas those based on research use complex statistical techniques that are still not fully understood. They give no indication of the relative importance of individual skills, nor do they provide guidance on how they should be sampled for test construction. Furthermore, the unspoken assumption is that these sub-skills described here are in fact skills – namely that they are things that somehow exist within the listener. There is reason to question whether this is the case. As we have seen, language use can be defined at almost any grain size – from broad sub-skills, such as the ability to listen for gist, to small detailed sub-skills, such as the ability to recognise the stress patterns of specific words. The research seems to suggest that we are able to identify them statistically at almost any level of detail. Perhaps these sub-skills are just ways of describing language-use activities. In other words, they may not be something we have within us, but may just be something we do with language. Nevertheless, what these taxonomies of sub-skills do is help us think about what processes we should be including in our listening tests.

Overall, what do these taxonomies tell us about the listening construct? The first point is that listening is a multi-faceted process, with a large number of sub-components that can be viewed from a number of different perspectives. The short taxonomies make the point that we need to test both the ability to extract the basic linguistic information, and the ability to interpret that in terms of some broader context. We need linguistic processing, which should include phonology, stress and intonation, word meanings, syntax and discourse features. We also need interpretation in terms of the co-text, the context of situation and world knowledge, which would include

summarising, making inferences, understanding sociolinguistic implications, following the rhetorical structure and understanding the speaker's communicative intent.

Conclusion

In Chapter 1 we looked generally at how comprehension works, whereas in this chapter we have concentrated on the characteristics unique to the listening situation.

We have seen how the sounds of the language often undergo considerable modification in rapid speech; and that there is considerable redundancy in the input, and much of the information is provided in a number of different forms. Stress and intonation are important and carry vital information. Stress is used to mark important words or ideas, and intonation is used to mark the grammatical structure of the text as well as indicating the attitude of the speaker to the content of the message.

Speech is usually a quick first draft, and often contains a variety of hesitations, repetitions and corrections. Spoken texts usually consist of a series of short utterances, strung together in a simple linear manner. Ideas may be developed over a number of utterances. The relations between ideas are often expressed less through grammatical structure but more through the juxtaposition and coherence of the ideas. Often it is not clear exactly what pronouns or cohesive markers refer to.

Spoken texts exist in real time and need to be processed very rapidly as they are heard. We have seen how the listeners' understanding of a passage depends very much on the purpose for listening and the background knowledge of the listener. Then, having been spoken, the text is gone, and generally all the listener has left is a memory of what they understood.

I have examined a number of ways of defining the listening construct in terms of its sub-components. In the next chapter, Chapter 3, I will look at how language testers have developed notions of the listening construct, and how they have attempted to use these to develop test items. Then, in Chapter 4, I will describe how to provide construct definitions useful for test development.