

2 What speakers know

- Extralinguistic knowledge
 - Sociocultural knowledge
- Linguistic knowledge
 - Genre knowledge
 - Discourse knowledge
 - Pragmatic knowledge
 - Grammar
 - Vocabulary
 - Phonology
- Speech conditions

In the last chapter we treated speaking as a **skill**, in the same way you might talk about the skill of playing the guitar or of driving a car. But being *skilful* assumes having some kind of knowledge base. To play a guitar well requires some kind of musical knowledge; to drive a car requires knowing something about how it works, as well as knowing the highway code. Of course, the knowledge base for speaking in a first language is largely intuitive; it is not something a person is normally sufficiently aware of to be able to describe. In order to describe it, researchers are compelled to infer it, both from the evidence of actual performance and also by studying the way it develops in early childhood. We shall draw on that evidence to identify what it is that speakers *know*. Knowledge that is relevant to speaking can be categorized either as knowledge of features of language (**linguistic knowledge**) or knowledge that is independent of language (**extralinguistic knowledge**).

Extralinguistic knowledge

The kinds of extralinguistic knowledge that affect speaking include such things as topic and cultural knowledge, knowledge of the context, and familiarity with the other speakers. In the dinner-party conversation about kedgeree (page 2), the speakers share considerable background knowledge at all these levels, and this is reflected in the assumptions they are able to make. Kath, for example, doesn't have to explain what *domestic science* is, and the other speakers throw in references to *spotted dick*, *toad-in-the-hole*, and *galub jalum*, as if these concepts were part of their common experience. These all constitute topic and cultural knowledge.

Context knowledge allows speakers to make reference to the immediate context, as in Hilda's mention of *this Rioja* (referring to the wine they are drinking). The use of mild oaths, such as *I mean for God's sake* (turn 63), suggests that, in this conversation, there is a level of interpersonal familiarity between the speakers that permits a degree of informality that would not be the case with total strangers.

Of course not all speaking events can rely on quite such a degree of shared knowledge. Explaining street directions to a total stranger or giving a lecture on quantum theory are both uses of speech that will require a considerable degree of explicitness. But because most speaking takes place face to face, and in a shared context, there is generally less need to be as explicit as one might normally be in writing, for example. After all, if your interlocutors don't understand you, they only have to ask. This 'situated' nature of speech means that it is characteristically **elliptic**: i.e. words, phrases, whole clauses are left out because they are redundant. So, when Hilda, in turn 60, says: *I would just make egg and bacon*, what is understood is the unstated idea: ... *when I did domestic science at school*.

Other characteristics of spoken language that derive from its being grounded in a shared context are:

- high frequency of personal pronouns, especially *you* and *I*;
- the use of substitute forms, as in (turn 59) *I had to sit there while everybody else did*, where *did* substitutes for *made kedgerree*;
- and the use of **deictic** language, that is, words or expressions that make direct reference to the context, as in *this Rioja*.

Sociocultural knowledge

'In X country long silences are tolerated in conversations.'

'In Y country you don't normally ask people why they are not married.'

'In Z country you always refuse an offer at least three times before accepting.'

Statements like these belong to the area of **sociocultural knowledge**. This is knowledge about social values and the norms of behaviour in a given society, including the way these values and norms are realized through language. Sociocultural knowledge can be both extralinguistic and linguistic. Knowing whether people in a given culture shake hands on meeting, or embrace, or bow, is extralinguistic; knowing what they say when they greet each other is clearly linguistic.

There has been a lot of debate as to the extent to which cultural differences cause misunderstandings or even breakdowns in communication. Unfortunately, the topic gives rise to a great deal of 'folk theorizing' and cultural stereotyping, of the type *All Japanese do such-and-such* and *All Arabs say so-and-so ...*. In fact, studies of conversational style suggest that there may be as many differences *within* a particular culture as there are between cultures. In any group of talkers anywhere, there is always someone who will dominate the conversation and someone else who won't say very much at all.

Nevertheless, there are certain speech events, such as greetings, requests, or apologies, where the risk of causing offence has meant that these events

have become ritualized in different ways across social groups. Part of a speaker's knowledge, then, is knowing what these sociocultural rules are and how they are codified.

Linguistic knowledge

Linguistic knowledge is often ranged along a cline from 'the big picture', e.g. knowledge of the way an anecdote typically unfolds, to the 'fine print', e.g. knowledge of grammar and vocabulary. In fact, the boundaries between categories are blurred, and they work interdependently, such that in reality it is difficult to account for particular features of a speech event by reference to any single knowledge system. However, for convenience, we shall discuss these different levels in turn.

Genre knowledge

Very broadly, there are two main purposes for speaking. Speaking serves either a **transactional** function, in that its primary purpose is to convey information and facilitate the exchange of goods or services, or it serves an **interpersonal** function, in that its primary purpose is to establish and maintain social relations. A typical transactional speech event might be phoning to book a table at a restaurant. A typical interpersonal speech event might be the conversation between friends that takes place *at* the restaurant. The story that Kath tells about her domestic science class is motivated less by the need to convey the facts of the matter (i.e. a transactional purpose) than by the wish to amuse her audience and thereby maintain a sense of shared community between friends (i.e. an interpersonal purpose).

These two basic purposes for speaking generate a host of different types of speech events. These, in turn, will be sequenced and structured in accordance with the kinds of social and mental processes that they accompany. We saw, for example, how Kath told her kedgeree story according to a narrative script, which, to put it very simply, has a beginning, middle, and end.

Service encounters, such as buying goods, getting information, or requesting a service, are transactional speech events that follow a fairly predictable script. Typically, the exchange begins with a greeting, followed by an offer, followed by a request, and so on, as in:

Good morning.
 Good morning.
 What would you like?
 A dozen eggs, please.
 Anything else? ...
 etc.

A certain amount of variation is generally permitted: some of the moves may be dispensable, while others of a more interpersonal nature – such as a comment about the weather – might be optional. Different cultures and sub-cultures may develop their own variants. Some service encounters in some cultures may permit bargaining, for example.

Over time and within particular speech communities, certain ways of realizing these speech events have become conventionalized to the point

that they have evolved into specific **genres**. Genre is an elusive term. Here we will use it to mean simply a type of speech event, especially in terms of how that speech event might be labelled by its participants. Hence, there is a difference between saying 'I had a chat with the boss' and 'I had a job interview with the boss' or 'I did a presentation to the boss'. Knowledge of how specific genres – such as chatting, job interviews, or business presentations – are realized is part of the linguistic knowledge that speakers in a particular speech community share. (How genres are integrated into genre-based teaching programmes is discussed in Chapter 7.)

An important factor that determines the structure of a genre is whether it is **interactive** or **non-interactive**. Multi-party speech, as in a shopping exchange or casual conversation between friends, is jointly constructed and interactive. Monologues, such as a television journalist's live report, a university lecture, or when you leave a voice-mail message, are non-interactive.

Finally, a distinction needs to be made between **planned** and **unplanned** speech. Certain speech genres, such as public speeches and business presentations, are typically planned, to the point that they might be completely scripted in advance. This means that their linguistic features will resemble or replicate features of written language. On the other hand, a phone conversation to ask for train timetable information, while following a predictable sequence, is normally not planned in advance: each participant has to make strategic and spontaneous decisions on the basis of the way the discourse unfolds. This, in turn, will affect the kind of language used.

On the basis of these criteria, we can classify speaking genres according to their general purposes, the kind of participation they involve, and the degree of planning (bearing in mind that these distinctions are less polarities than stages on a continuum). For example:

	purpose	participation	planning
airport announcements	transactional	non-interactive	planned
sports commentary	transactional	non-interactive	unplanned
job interview	transactional	interactive	(partly) planned
service encounter	transactional	interactive	unplanned
joke telling	interpersonal	(partly) interactive	(partly) planned
leaving a voice-mail message	transactional or interpersonal	non-interactive	unplanned
casual conversation	interpersonal	interactive	unplanned

Discourse knowledge

Within the structure of a specific genre, its individual elements need to be connected so as to form coherent stretches of discourse. Knowing how to organize and connect individual utterances, as well as how to map this

knowledge on to the turn-taking structures of interactive talk, is called **discourse competence**. For example, when Kath says:

It's one of those ridiculously old-fashioned dishes that they make you cook in domestic science

Nick responds:

Well why don't you try making some? Might be great

His use of the **discourse marker** *Well* serves to link his utterance to Kath's previous turn, while the pronoun *some* substitutes for the previously mentioned *kedgerie*, referred to as *it* by Kath. Likewise, the ellipted *it* in Nick's utterance *might be great* also refers back to *kedgerie*. Further cohesion between the two turns is achieved through the use of the synonyms: Kath's *cook* is echoed in Nick's use of *making*. Thus, the speakers are drawing on their lexical and grammatical knowledge to make connections between utterances and across turns, within the strict constraints of the rules of turn-taking.

The use of discourse markers is particularly important in terms of the fluid management of interactive talk. Discourse markers are used to signal one's intentions, to hold the conversational turn, and to mark boundaries in the talk. For example, in the following extract Kath signals, in turn 61, that she hasn't quite relinquished the topic of *kedgerie*, nor drawn a moral for her story, despite Hilda's comment about egg and bacon.

(60) Hilda: I would just make egg and bacon
(61) Kath: But *kedgerie*. This was a sort of comprehensive school the first year of.

Kath uses the discourse marker *But* to retrieve the topic, to connect her utterance with her previous story, and to signal the contrastive nature of the conclusion she wants to draw.

Here are some common discourse markers and their meanings:

- *right, now, anyway*: these mark the beginning or closing of a segment of talk.
- *well*: this is a very common way of initiating a turn and linking it to the preceding turn, often to mark the onset of a contrast, e.g. a difference of opinion.
- *oh*: this is typically used either to launch an utterance or to respond to the previous speaker's utterance, often with implications of surprise or unexpectedness.
- *and, but, or*: these conjunctions are used to connect discourse: *and* marks some kind of continuity, *but* marks a contrast, and *or* marks an option.
- *so, because*: these are also conjunctions: they signal that what follows is (respectively) the *result* or the *cause* of what has been mentioned.
- *then*: this is often used to signal an inference based on what someone else has said.

- *y'know, I mean*: these markers serve to gain and maintain attention on the speaker – the first by appealing to the hearer's shared knowledge, and the second by signalling that some kind of clarification is going to follow.

Pragmatic knowledge

Pragmatics describes the relation between language and its contexts of use, including the purposes for which language is being used. How do speakers adjust their message to take context into account? And how do listeners use contextual information to make sense of what they are hearing?

- **Speech acts**

A communicative view of language holds as axiomatic that when someone says something, they are also *doing* something. For example, in the kedgere conversation, Hilda (turn 53) says:

This is really nice this Rioja

and at almost the same time Nick (turn 54) says:

Well why don't you try making some [kedgere]? Might be great

Both Hilda's and Nick's utterances have a communicative purpose: Hilda's utterance functions as praise; Nick's as a suggestion. There are both lexical and grammatical clues that help us in assigning a function to these utterances. For example, the structure *This is really X, this Y* is a very common way of making an evaluation in spoken language. Likewise, *Why don't you ... ?* is a common way of framing a suggestion.

The way that specific **speech acts** (also called **functions**), such as complementing, suggesting, requesting, offering, and so on, are typically realized comprises part of a speaker's **pragmatic knowledge**. Pragmatic knowledge is knowing how to do things with language, taking into account its contexts of use. This, in turn, means knowing how to perform and interpret specific speech acts. Knowing that one way of framing a request is *Would you mind if ...*, as in *Would you mind if I turn the volume down?*, is part of pragmatic knowledge. It is also knowing that speech acts can be realized indirectly – that, for example, the statement *the music is very loud* has the force of a request (to turn the music down), if uttered in certain contexts.

Because speech acts often have an instrumental function, in that they involve getting people to do things, they typically form one part of a reciprocal exchange. For example, it is normal to respond to a request with some kind of agreement:

Would you mind if I turn the volume down?
Not at all.

Paired utterances like this, in which the second is dependent on the first, are called **adjacency pairs**. Questions and answers are the most common form of adjacency pair, as in:

Simon: Have you ever eaten kedgeree since?
Kath: Oh yes I love kedgeree.

But also greetings, requests, invitations and offers, compliments, reprimands, and apologies are all exchanges that are typically realized by means of adjacency pairs. Often, too, they are quite formulaic, as in the case of greetings:

How do you do?
How do you do?

In fact, many so-called pairs have a three-part structure, where the first speaker adds some kind of evaluation:

Would you mind if I turn the volume down?
Not at all.
Thanks.

Three-part exchanges are very common in classroom talk:

Teacher: What's the past of the verb to go?
Student: Went.
Teacher: Good.

This three-part instructional sequence is called an **IRF exchange**, IRF standing for **initiate – respond – follow-up**.

Longer sequences of paired utterances are also a feature of the *openings* and *closings* of conversations. Take, for example, this closing of a telephone conversation:

Well, I'd better get back to work. Hmm, me too. So, I'll speak to you later. OK, then.	<i>pre-closing</i>
Have a good day. You, too. Bye. Bye bye.	<i>closing</i>

Speech act knowledge, then, means knowing not just how particular speech acts are typically realized, but how such speech acts fit into the longer exchanges that form units of talk.

- **The co-operative principle**

Interpreting the communicative force of speech acts, and knowing how to respond appropriately, assumes that participants in a speech event are 'playing the game according to the same rules'. For example, if you ask a question, you assume that what your interlocutor says in response is an answer. Or, if not, that it is nevertheless relevant to what you have just

asked. For example, in this exchange, Bea's response to Andy's question is also a question:

Andy: What does *pragmatics* mean?
Bea: Do you have an hour or two?

Because Andy takes for granted that Bea is co-operating in the conversation, he has to assume that she isn't ignoring his question and initiating another conversational topic altogether, but that her question is somehow relevant to his question. So he says:

Andy: Complicated, huh? Just give me the short answer.

Andy correctly understood that Bea's answer implied that pragmatics is a complicated subject, not amenable to a snappy definition. The assumption that, in the absence of any evidence to the contrary, speakers are co-operating with one another forms what is called the **co-operative principle**, which the philosopher Grice elaborated into four maxims:

- 1 Quantity: Make your contribution just as informative as required.
- 2 Quality: Make your contribution one that is true.
- 3 Relation: Make your contribution relevant.
- 4 Manner: Avoid obscurity and ambiguity. Be brief and orderly.

Sometimes it is not easy to abide by these maxims, and speakers will often indicate that they may be at risk of violating one of them. For example, in this continuation to Andy and Bea's conversation, Bea starts by saying:

Well, *at the risk of oversimplifying matters*, pragmatics is about language in context ...

She is aware that her answer may be less informative than necessary, thereby running counter to the maxim of quantity. On the other hand, she might have said:

Well, *I may be wrong but I think* pragmatics is about language in context ...

which would indicate that she is aware that her answer may not be accurate, hence a potential violation of the maxim of quality. The frequency of such hedges is a good indication of the extent to which speakers are aware of the rules underlying the joint construction of meaningful talk.

- **Politeness**

The rules of conversational co-operation should not be confused with **politeness**. In fact, given the choice between saying the truth and not hurting someone's feelings, speakers will usually opt for the latter, as in this instance:

Bea: What did you think of my presentation?
Andy: I thought it was very well researched.

Because the research is only one aspect of the presentation, Andy is not really making his contribution as informative as required (thus he is flouting the maxim of quantity). He does this in order not to threaten Bea's face, that is, her social standing and sense of self-worth. Contrast his 'faint praise' with a more direct affront to face, such as *I didn't like it very much*. Politeness, then, refers to the way we take other speakers' face needs into account.

Languages employ an elaborate armoury of means to avoid threats to face. The use of politeness markers, such as *please* and *thank you* (or their equivalents), are universal. In some languages, positive politeness is encoded in the pronoun system. In French, for example, speakers can choose between *tu* and *vous*, according to the degree of familiarity or respect they wish to convey. In English, as in many languages, the use of distancing devices, such as past tense forms and modal verbs, helps soften the potential threat to face of requests or commands:

I was wondering if you were free on Friday.
Could you turn the lights out when you leave?

Knowledge of how politeness is encoded in the language is obviously a crucial component of knowing how to speak.

- **Register**

Politeness requires of speakers a sensitivity to context, especially the **tenor** of the context – that is, the relationship between speakers, including such factors as relative status and familiarity. Other factors in the context of the speech event will also impact on the language used, particularly on its degree of formality. (Note that formality and politeness intersect, but that they are not the same thing: you can be formal and rude, just as you can be informal and polite.)

Along with tenor, the linguist Michael Halliday identified two other key dimensions of context: the **field** and the **mode**. The field of a speech event refers to the *what* of the event – what is going on, what is being talked about, such as 'a lecture on nutrition', or 'a conversation about food'. The tenor, as we have seen, refers to the *who*, and the mode refers to the *how* – the choice of channel, such as whether the speech event is conducted over the phone as opposed to face-to-face, or in real-time as opposed to prerecorded. Together, these three contextual factors – field, tenor, and mode – influence the speaker's choice of **register**, such as where the speech event lies on a continuum from formal to informal, and whether it is characterized by jargon and other in-group language forms. The register of a university lecture on the topic of nutrition will differ markedly from a conversation between friends on the subject of domestic science. You wouldn't expect to hear the university lecturer say: *Kedgeriee, I remember saying to my mum, I've got to take a pound of fish next week we're making kedgeriee ...* . In fact, even the term *mum* would sound out of place.

A speaker's knowledge, then, involves knowing what language choices are appropriate, given the register variables of field, tenor, and mode.

Grammar

It is theoretically possible to have short conversations where each utterance consists of nothing but a single word or short phrase, as in this invented example:

A: Coffee?
 B: Thanks.
 A: Milk?
 B: Please.
 A: Sugar?
 B: No, thanks.

In this instance, context factors, including the lack of formality, make the use of complex language unnecessary. But to sustain a conversation like this over a variety of topics with a number of speakers would be virtually impossible. The effect would be like baby talk. In order to generate a much more sophisticated range of meanings, the resources of the language's grammar need to be enlisted.

This does not mean, however, that the grammar of speech is identical to the grammar of written texts. We have already noted how the demands of producing speech in real-time with minimal planning opportunities places considerable constraints on the kind of complexity speakers can achieve. A sentence like that last one is much more typical of written language than of spontaneous spoken language. Spoken, it might have sounded like this:

Speaking, you're doing it in real-time, you don't have much planning time, so it tends to be less complex than ... or rather it's a different kind of complexity, than, say, writing.

Another distinguishing feature of spoken grammar is the three-part division of utterances into a body plus optional head and tail slots, as in:

head	body	tail
Kedgerree	I remember saying to my mum ...	
	This is really nice	this Rioja

Not to be confused with tails are **tags**, typically question tags, with which the speaker makes a direct appeal for the listener's agreement, consent, and so on. They therefore have a primarily interpersonal function. For example:

body	tail	tag
This is really nice	this Rioja	isn't it?

Question tags are virtually non-existent in written language, apart from in fiction, but they are extremely common in speaking, comprising a quarter of all questions. Other ways of forming a tag include expressions like *right? no? ok?* and the vernacular *innit?*

Other features of spoken grammar that are less rules than tendencies are a preference for direct speech rather than reported speech, as in:

she said 'you don't want to be making kedgeree' and she said 'we don't like it'

and the use of **vague language**, as in:

It's a sort of old colonial dish

Vagueness expressions are used not only to fill pauses, but also to reduce the assertiveness of statements. This is a way of fulfilling Grice's 'maxim of quality' (*make your contribution one that is true*). It is also a way of reducing the face-threatening potential of an assertion – of being less 'bold'. Writing, however, typically requires greater precision, or may use other means, such as modality, to reduce the assertiveness of statements (as in this sentence).

We have also seen how spoken language tolerates **ellipsis**, as in *Might be great*, where in writing *It might be great* would normally be preferred.

Finally, there are a number of features of spoken grammar that are the audible effects of real-time processing difficulties – what we will call **performance effects**. These include the use of hesitations (*erm, uh*), repeats, false starts, incomplete utterances, and **syntactic blends**, i.e. utterances that 'blend' two grammatical structures, as in *I've been to China ... in 1998*.

Features of spoken grammar that distinguish it from written grammar are summarized in this table:

Written grammar	Spoken grammar
Sentence is the basic unit of construction	Clause is the basic unit of construction
Clauses are often embedded (subordination)	Clauses are usually added (co-ordination)
Subject + verb + object construction	Head + body + tail construction
Reported speech favoured	Direct speech favoured
Precision favoured	Vagueness tolerated
Little ellipsis	A lot of ellipsis
No question tags	Many question tags
No performance effects	Performance effects, including: <ul style="list-style-type: none"> • hesitations • repeats • false starts • incompleteness • syntactic blends

Other differences between written and spoken grammar have to do with the distribution of particular items. We noted, for example, that personal pronouns and determiners (such as *I, you, my, our ...*) are more frequent in

spoken language than they are in written. The following list summarizes facts about the distribution and frequency of verb forms in spoken language:

- present tense forms outnumber past tense forms by 2:1.
- simple forms outnumber progressive and perfect forms by over 10:1.
- the past perfect and present perfect continuous are rare.
- passive verbs account for only 2% of all finite verb forms in speech.
- *will*, *would*, and *can* are extremely common in speech.

Vocabulary

The Russian theorist Bakhtin hypothesized a ‘fully meaningful and complete’ conversation between two people in a room that consisted of nothing but the one word: *Well!* In fact, a lot of conversation does consist to a very large extent of such common words and short phrases as *well*, *yeah*, *but*, *I know* etc. Researchers, using large databases (**corpora**) of transcribed speech, have demonstrated that the fifty most frequent words in spoken English make up nearly 50% of all talk. (This contrasts with a figure of less than 40% of coverage for the fifty most frequent words in written English.) As an example, the word *well* occurs about nine times more often in speech than in writing.

Well is an example of a discourse marker (see above) which is very common in spoken interaction. Spoken language also has a relatively high proportion of words and expressions that express the speaker’s attitude (**stance**) to what is being said. These include ways of expressing doubt and certainty, such as *probably* and *maybe*, as well as ways of emphasizing the factual nature of what is being said, such as *really* and *actually*.

Speakers also employ a lot of words and expressions that express positive or negative **appraisal**. This is due to the fact that a lot of speech has an interpersonal function, and, by identifying what it is they like or don’t like, speakers are able to express solidarity with one another. In this short extract from the kedgerie conversation, the appraisal language is underlined:

- (52) Kath: It’s one of those ridiculously old-fashioned dishes that they make you cook in domestic science =
- (53) Hilda: This is really nice this Rioja
- (54) Nick: Well why don’t you try making some? Might be great

Finally, we have already mentioned the prominent use of **deictic** language in speech – that is, words and expressions that ‘point’ to the place, time, and participants in the immediate or a more distant context. The exact referents of deictic expressions – that is, the exact things or people they refer to – are only recoverable by reference to the context. Here are some common deictic expressions:

spatial deixis	temporal deixis	person deixis
here, this (place, thing etc)	now, this (time)	I, me
there, that (place, thing etc)	then, that (time)	you, your

So far we have talked about the types of words that are common in speech, but we haven’t said anything about the number. How many words do speakers

know? Here we need to distinguish between the words that speakers use (their **productive** vocabulary) and the words that they recognize (their **receptive** vocabulary). Research suggests that the former is only half the size of the latter. And the number of words used in speaking is less than the number used in writing. That is to say, in speech fewer words go further. According to some estimates, a vocabulary of just 2,500 words covers nearly 95% of spoken text (compared to 80% of written text).

- **Chunks**

As we saw in the discussion of Kath's kedgeroo story in Chapter 1, speakers achieve fluency through the use of prefabricated chunks. These are sequences of speech that are not assembled word by word but have been pre-assembled through repeated use and are now retrievable as single units. Chunks can be defined very broadly as any combination of words which occur together with more than random frequency. They are also known as **lexical phrases**, **holophrases**, **formulaic language**, and 'prefabs'. Of the different types of chunk, the following are the most common:

- **collocations** – such as *densely populated*, *rich and famous*, *set the table*
- **phrasal verbs** – such as *get up*, *log on*, *run out of*, *go on about*
- **idioms, catchphrases and sayings** – such as *part and parcel*, *make ends meet*, *as cool as a cucumber*, *speak of the devil*
- **sentence frames**, i.e. the fixed components of sentences, especially at the beginnings of sentences, that 'frame' open slots – such as *would you like a ... ? the thing is ...*, *what really gets me is ...*
- **social formulas** – such as *see you later*, *have a nice day*, *mind your head*
- **discourse markers** – such as *if you ask me*, *by the way*, *I take your point*, *to cut a long story short ...*

In the following short conversational extract between two Australian speakers of English, the likely chunks have been underlined (likely, because, without a more extensive study of each speaker's language, it is not easy to determine what is prefabricated as opposed to what is a novel construction):

```

<S 02> They were awake at five-thirty Stefan.
<S 01> Yeah.
<S 02> It's way too early after a night like they had last night.
<S 01> Yeah that's right. ... Yeah. You live and learn.
<S 02> And sometimes you don't live and learn. Sometimes you live and
repeat and repeat and repeat.
<S 01> [chuckles] Would you like a cup of tea?
<S 02> No thanks. I'm sick of that alarm going off all the time.
<S 01> Yeah.
<S 02> Don't you?
<S 01> Yeah. What can you do?

```

Even from this short extract, it's clear that chunks make up a large proportion of spoken language.

Some researchers estimate that a native speaker may have hundreds of thousands of these chunks to draw on, and that this accounts for both fluency (as we saw with the race-caller on page 7) but also for **idiomaticity**. By idiomaticity is meant the fact that, out of all the many possible grammatically acceptable ways of expressing an idea, speakers of a speech community tend to conform to what other speakers do. For example, it makes perfect sense, and it is grammatically correct, to say *it's six less twenty*, or *it's forty past five*, or *it's two thirds past five*, yet the 'done' way of expressing this idea is *it's twenty to six*.

The lexical knowledge that a proficient speaker has access to, then, consists not just of a few thousand words, but of a much greater number of chunks. Language corpora are starting to provide information as to which of these chunks are the most frequently used. For example, in a recent study of a corpus of spoken US English idioms, one researcher listed the following as some of the most frequent: *kind of, sort of, of course, in terms of, in fact, deal with, at all, as well, make sure, go through, first of all, in other words*.

Phonology

The 'lowest level' of knowledge a speaker draws on is that of pronunciation. Normally, the way we pronounce individual words, and the sounds that they are composed of, is not something that involves conscious choices. Words are stored along with their pronunciation and do not need to be reconstituted from scratch each time they are used. Occasionally, however, speakers will adjust their pronunciation to take account of the social context, so as not to sound too 'posh', for example. Or they will adopt an accent or a quality of voice for a particular dramatic effect. When, for example, Kath told her kedgeriee story, she adopted – and exaggerated – her mother's accent, to the amusement of her friends.

One area of pronunciation, however, where significant choices are available to speakers is in **intonation**. Intonation serves both to separate the stream of speech into blocks of information (called **tone units**) and to mark information within these units as being significant. In English, there is a fundamental association between high pitch and new information. So, within each tone unit, information that is being added to the discourse is made prominent through the use of a step up in pitch. Intonation also serves to signal the connections between tone units. Typically, a rise in pitch at the end of the tone unit (that is, after the last stressed word) implies some kind of continuation; a fall in pitch suggests completion.

In the following extract from Kath's story, the three functions of intonation (segmentation, prominence, and cohesion) are shown working in unison. The tone units are marked with vertical lines, and the words (or the parts of them) that are given prominence through a rise in pitch are capitalized. Pitch changes at the boundaries of tone units are marked by rising or falling arrows:

KEDgeriee ↗ | I reMEMber saying to my MUM ↗ | I've got to take a POUND of FISH next week ↗ | we're making KEDgeriee ↘ | and SHE said ↗ | you don't want to be making KEDgeriee ↗ | and she SAID ↗ | we don't LIKE it ↘ |

A further point to note here is the use of a marked rise in pitch on the first word of the story (*Kedgeree*), separating it from the preceding and surrounding discourse. This use of intonation to mark the beginning of a new stage in the discourse – equivalent to starting a new paragraph in writing – is called a **paratone**. It is very perceptible when news readers, for example, move from one story to the next, and it is balanced by an equally marked drop in pitch at the end of each story. Likewise, Kath's closing comment on her anecdote ends on a 'dying fall':

it was so inappropriate for the first year comprehensive school kids to be making ↘

And, if you were reading this paragraph aloud, you would also no doubt finish on a falling paratone.

Speech conditions

Kath was able to tell her story fluently because she knew it, she had told it before, and she was among friends. This suggests that the conditions in which speaking occurs play a crucial role in determining the degree of fluency that is achievable. What are these conditions? That is, what factors make speaking easy or difficult? Researchers have isolated a number of factors, of which the following seem to be the most important. They have been divided into three categories: **cognitive** factors, **affective** (that is, emotional) factors, and **performance** factors.

Cognitive factors

- *Familiarity with the topic*: the greater the familiarity, the easier the speaking task; this is why it is generally easier to talk about your job, or your family, than it is to talk about something very removed from your day-to-day life.
- *Familiarity with the genre*: giving a lecture or a speech will be harder if you're unfamiliar with those particular genres.
- *Familiarity with the interlocutors*: generally speaking, the better you know the people you are talking to and the more shared knowledge you can assume, the easier it will be.
- *Processing demands*: if the speech event involves complex mental processing, such as that involved in describing a complicated procedure without recourse to illustrations, it will be more difficult than if not.

Affective factors

- *Feelings towards the topic and/or the participants*: generally, if you are well disposed to the topic you are talking about, and/or to the other participants, the easier it is likely to be.
- *Self-consciousness*: being 'put on the spot' can cause anxiety which will have a negative effect on performance; likewise, knowing (or believing) that you are being evaluated can be prejudicial.

Performance factors

- *Mode*: speaking face-to-face, where you can closely monitor your interlocutor's responses and where you can use gesture and eye-contact, is

- generally easier than speaking over the telephone, for example.
- *Degree of collaboration:* giving a presentation on your own is generally harder than doing it with colleagues because in the former case you can't count on peer support.
 - *Discourse control:* on the other hand, it is often easier if you can control the direction of events, rather than being subject to someone else's control.
 - *Planning and rehearsal time:* generally, the more time to prepare, the easier the task will be; telling a joke is usually easier the second time round.
 - *Time pressure:* if there is a degree of urgency, it is likely to increase the difficulty for the speaker.
 - *Environmental conditions:* trying to speak against a background of loud music or in poor acoustic conditions (as in many classrooms!) is difficult.

The above factors do not necessarily predict the difficulty or ease of speaking since they also interact with personality factors, such as introversion and extroversion. It is not always the case, for example, that being put on the spot, or urgency, can have negative effects: some speakers respond positively to such pressure. Likewise, physiological factors such as tiredness can undermine performance. Nevertheless, the above factors offer a useful template for predicting the degree of fluency speakers are likely to achieve. (And, as we shall see in Chapter 6, they provide criteria for the selecting and adapting of classroom speaking tasks.)

Conclusions We started this chapter by making a distinction between what speakers *can do* – that is the mental and physiological processes involved in speaking – and what speakers *know* – that is the knowledge base that speakers draw on that enables these processes.

The kinds of knowledge that speakers bring to the skill of speaking comprise extralinguistic knowledge, such as background knowledge of topic and culture, and linguistic knowledge, including discourse knowledge, speech act knowledge, and knowledge of grammar, vocabulary, and phonology.

Looking ahead So far we have described speaking skills and speaker knowledge insofar as they relate to highly-skilled, knowledgeable speakers, making no distinction between speaking in a first or a second (or third, or fourth etc) language. But speakers of another language do not, initially, have easy access to these skills and this knowledge. In the next chapter, we will look at the implications of this skills and knowledge gap, and discuss general approaches to how it might be bridged.