Chapter 2

The research space: paradigms and issues

This chapter will...

• describe some classical research paradigms as they relate to spoken mode and to analysing spoken data;
• discuss particular issues surrounding research into speaking;
• discuss the role and status of spoken data in language theory.

2.1 Introduction

This chapter begins with a brief introduction to research paradigms in general, and then discusses their usefulness and applicability in relation to researching speaking. In this chapter, I also address the question of the nature of spoken data and how this relates to the kind of research that is undertaken into spoken mode.

2.2 Classical research paradigms in relation to researching speaking

The research approaches dealt with in this book are, in general, empirically based. That is to say, they deal in real-world data of some kind – systematically recorded observations of classroom behaviour, transcripts of conversation, recordings of learners’ utterances analysed for the occurrence of particular phonemes and so on. These data are gathered to investigate a central research question, often posed as a hypothesis, and are used as the
basis of either a quantitative analysis (most often) or a qualitative analysis (less frequent apart from areas such as critical linguistics, socio-linguistics and ethnographically based work).

A second type of approach is a more theoretically oriented one. Here, rather than taking data as the starting point of an investigation, the researcher is primarily interested in theories, models, high-level concepts, and, crucially, the relationships between previous theories and ones that may emerge from a current investigation. In the context of extremely theoretically oriented work, for example philosophical logic (a discipline with a surprising amount of influence on linguistics), any real-world data are, if they are considered at all, seen as ‘messy’, subject to the vagaries of individual circumstances and irrelevant.

Different research approaches are often called ‘paradigms’ (see Concept 2.1) and these strongly influence how research is carried out. A paradigm functions as a framework or point of reference for both researchers and users of research output. It gives coherence to a study and links it to the work of others providing a shorthand by which it can be prepared and judged. If the researcher positions a study in, for example, an experimental paradigm he or she creates a set of expectations in his or her audience about the way the research will be conducted. These considerations affect how research is received.

Concept 2.1 What is a paradigm?

A paradigm is a framework for ideas which includes definitions of key terms and the relationships between them. The framework is coherent because the researcher assumes certain things as a starting point and new knowledge is absorbed into this mental ‘map’. As noted in the previous chapter, in the US Noam Chomsky created a major shift in what people assumed about language when he conceptualised it as an inherent rule-governed system for which the human mind is hard-wired from birth, and also set up key concepts such as performance versus competence. This was in sharp contrast to pre-existing paradigms in American linguistics which had been strongly data-oriented/ethnographic in nature, and also contrasted to the European structuralist paradigms that had emerged after the work of Ferdinand de Saussure. This kind of change is referred to as a ‘paradigm shift’.

Different disciplines work within different paradigms and even within the same academic department several paradigms can compete with one another. Most research outcomes make only small changes to the paradigm rather than altering it fundamentally – this is the nature of research findings generally. Paradigm shifts can and do occur when either a brilliant individual or a team compel others to change their mental map of a particular topic due to the strength of their findings or arguments.
Concept 2.2  **Empirical versus non-empirical approaches**

An empirical approach to a research project begins from situated questions and facts rather than decontextualised issues or questions of theory. Both kinds of approach can be used to deal with the same topic and the clearest differences between them lie in the methods used and what is regarded as a coherent approach and evidence. The less empirically motivated researcher in linguistics has traditionally been concerned with intuitions about language and the fit between new data and existing theory. The empirically oriented researcher will be more open to seeing patterns emerging from data and drawing conclusions from these that may challenge pre-existing ideas or intuitions. Both approaches will draw on theory and a pre-existing paradigm so the contrast is not so much between empirical and theoretical work but between the emphasis placed on data in terms of conclusions reached. All research needs to abstract away from particular instances of data in order to reach some coherent conclusion but empirically grounded work will retain stronger links to concrete examples and give these greater weight than non-empirically oriented research.

All paradigms orient towards a theory and towards data, but the balance between these will differ according to the tradition in which the academic is working. Academic research is meaningless if it is not embedded in the context of the work of others. This work is in turn framed within a paradigm that has a particular orientation towards data and theory. Different disciplines will also place different emphasis on the role of theory versus data. In linguistics, and particularly in the realm of spoken discourse, the relationship is quite complex and the locus of ongoing debate. At one end of the spectrum, an academic working in the field of syntax will aim to achieve an elegant, comprehensive and convincing description of a language feature such as negation in a particular language and relate this to current theories of negation generally. To be convincing the work will need to orient towards all previous work on negation and will tend to do this within a theoretically oriented paradigm. While examples will be used, the work will rarely be ‘data driven’ in the way that the work of a text or corpus linguist will be. At the other end of the spectrum, in computational linguistics there are academics developing models of grammar via automatic ‘parsers’ purely from massive numbers of examples in ways that allow syntactic categories and patterns in language to be described in a bottom-up fashion.

Table 2.1 gives an indication of how far, in general, some of the major branches of linguistics deal with situated data, whether they regard mode as relevant, or in contrast deal primarily with abstractions.
The groupings shown in Table 2.1 should be regarded as broadly indicative and they are there to help situate a debate about the role of spoken data in language theory and, in turn, address the question of why there are few holistic theories of speech available. As an example, socio-linguistic research frequently draws its data from the spoken mode (for example, on the social marking carried by a particular phoneme or the speaking strategies of a particular racial group), but does not relate the findings to any broader theory of speech. It is important therefore to distinguish between research into speaking and research that uses speech data for a different research purpose.

There are reasons for the different status of data, and particularly speech data, in various branches of language studies. Linguistics as we know it today has a surprisingly short history and since the 1960s has been developing and positioning itself among several disciplines, newer and older than itself. In the early part of the twentieth century, what we call linguistics was termed the ‘science of language’. It was primarily interested in concrete examples of language, and the study of the history of the development of a language or the comparison of different languages (philology and its branches) were the focus of its efforts. There was then a transition from what was a largely descriptive analytical discipline (and one that in its attention to detailed contrasts and taxonomies was akin to botanical science and related disciplines) to one that set great store on the need to theorise...
away from the messy, real-world data, to universal regularities or competencies. This process has led to particularly interesting and complex issues surrounding the attitude to speech data in language theory generally, and the next sections will deal with this further.

Concept 2.3  The powerful influence of a compelling and coherent theory

An interesting example to flesh out the differences between empirically oriented work and those that take theory as a starting point is research on rhythm in spoken language. The classical paradigm set up and developed by among others Pike (1945) and Abercrombie (1967) proposes that languages should be categorised in terms of two different rhythmic systems: syllable timing and stress timing. In the former kind of language every syllable has the same duration and in the latter syllable length varies so that a regular ‘beat’ is created by the words and phrases of the language. Spanish and French are, traditionally, categorised as syllable timed and English and Russian as stress timed languages. This very compelling idea of a binary contrast (nicely described as ‘machine gun’ (syllable timing) versus ‘morse code’ (stress timing) (Lloyd James, 1940)) has held sway with variations for nearly 70 years. This is despite the fact that researchers admit that when they measure and time samples of languages it is difficult to find data that consistently fit the theory. Very complex systems of metrics have been created (Low and Grabe (1995), Grabe and Low (2002)) to investigate speech rhythm, most of which begin from this binary contrast or refer back to it. More recently the idea of stress/syllable timing being less clear-cut categories towards which individual languages tend, rather than being their defining rhythmic characteristics, has emerged but the paradigm remains largely unshifted or at least still has currency.

2.3  Attitudes to speech data

Quote 2.1  Attitudes towards speech data in linguistics in the early 1980s

Methodologically, most contemporary linguists do not use actual speech as a source of data for the analysis of linguistic structure. They base this position in part on the argument that the phrasal breaks, such as restarts, found in actual speech give evidence of such defective performance that the data are useless for the study of competence.

(Goodwin, 1981: 12)
Even theoretically oriented work engages with data at some level. At its most basic the research is grounded in some real-world concepts, if not ‘hard’ data. When researchers think of empirical approaches in opposition to more theoretically oriented ones, it is a matter of what role the data are seen to have in the research process. In ‘classical’ theoretically oriented, scientific methods, the model or theory on which a study is based is not going to be fundamentally redefined by the outcomes of the research. Data which challenge the prevailing theory are likely to be set aside as ‘blips’ and more generally the phenomena being investigated will be selected in such a way that they will tend to fit in with the existing paradigm (see Concept 2.3 for an example of this).

These are particularly pressing issues for the researcher into speech for three reasons. First, unlike the written form, the building blocks of speech do not come to us in a clearly demarcated set of units. Our literate view of language means that it is a surprise to realise that the stream of speech is exactly that: there are no gaps between individual words. The process of understanding speech is highly dependent on an interpretive capacity on the part of the listener and this interpretive role is not one that the researcher can completely stand apart from when handling authentic data. Second, capturing and analysing speech depends largely on the written form and careful attention is needed to the relationship between the original data and its visual representation – the secondary data. Finally, as noted above, neat and clearly defined categories and patterns are extremely compelling and there can be a tendency to ‘retrofit’ speech data to pre-designated categories due to this. Research into spoken grammar shows this particularly clearly. The terminology of traditional pedagogic or prescriptive grammars struggles to describe the norms of the spoken mode.
Concept 2.4 Finding words to describe the grammar of speech

Traditional and/or pedagogic grammar provides a fairly consistent set of constructs, definitions and structural relations. A grammatical construct like ‘relative clause’ or ‘noun phrase’ is relatively stable and clearly defined – a researcher will find several hundred articles on the topics with ease.

Research into the grammar of spoken discourse has suggested that there are a number of constructions regularly used by speakers (for example, subject-verb ellipsis – ‘Nice day’ as opposed to ‘It is a nice day’ (Nariyama, 2004)) which do not fit into the norms of traditional grammar models, or items which have a high occurrence (for example, semi-modal verbs such as ‘tend to’ (Moore, 2007)) but which are presented as ‘unusual’ in standard grammars. Structures such as these that fall outside the standard definitions are less easy to handle for two reasons. First, by their nature they do not fall into the neat categories of the existing grammar model. Second, there will be no accepted terminology for the elements being described. Thus, a construction typical of spoken English such as the following, ‘where he went wrong my mate Tony was not getting the car taxed before he went on his holiday’, might be defined as a ‘left’ sentence, ‘pre-posed’, containing a ‘left-shifted head’ or other terms which may or may not mean exactly the same thing to everyone or overlap with one another exactly.

In the first part of the twentieth century, speech itself was difficult to capture, and even the advent of the tape recorder meant that gathering large samples of data and analysing them was a laborious process. The ability to record speech, and the comparatively recent growth in the power of the personal computer, has brought the possibility of large corpus studies to the office of the applied linguistics researcher. However, the complexities of capturing large quantities of spontaneous spoken data have meant that most corpora still depend for their input on the written mode. Insights from corpora that combine a balance of both spoken and written material are beginning to filter into the public domain in forms that can be used by the teaching community. See for instance Biber et al. (1999) Longman Grammar of Spoken and Written English or Carter and McCarthy (2006) Cambridge Grammar of English.

It is noticeable, however, that despite advances in the capturing and the analysis of speech data, research questions continue to be oriented towards areas other than finding out more about the nature of speech, per se. Considering the universality of the ability to speak across humankind there has been little attempt to draw together a unified theory of the process. Many disciplines value real speech data and place them at the heart of their theories. However, these approaches have tended to incorporate the spoken language into a theory that aims to describe or explain
something else. For instance, second language acquisition (SLA) gives high importance to the effect of spoken input on the learner but the elements under discussion have tended to be the learner’s inherent capacity for language learning, the closeness or distance between a target language and current utterances, how their first language affects their second, and so on.

Notable exceptions such as Levelt’s seminal work *Speaking: from intention to articulation* (Levelt, 1989) fall outside what is considered core work in applied linguistics, coming under the umbrella of psycholinguistics. Even here work stops largely at the point of utterance and does not pursue the important issues of interaction, the influence of intonation and prosody, turn-taking and so on; nor how these features might relate to one another in a process of communication that is unique to the spoken mode. In language acquisition, research with an interest in bridging some of these gaps began to emerge in the early years of this century (see for instance, Judit Komos’ readable *Speech Production and Second Language Acquisition* (Komos, 2006)).

### 2.4 The applicability of research approaches and frameworks to the study of speech

The previous sections have argued that care is needed in researching speaking due in three respects. These were the strong influence of a literate view of the form, the tendency to tidy speech data and to abstract away
from the messiness of real-world, situated, talk in context, and the tendency to use speech data as the basis for research into some aspect of language other than the spoken mode in its own right. Here we look at the implications of these points, and what research into speaking _per se_ may, in due course, emerge as.

Hand in hand with a removal of the object of study to the theoretical, unsituated, or abstract level is a convenient merging of the construct ‘speech’ with ‘language’. It is convenient because it permits the models in question to use isolated examples closer to the norms of formal, published written mode and ignore deviant, ill-formed and difficult to parse forms which might come under debate if real-world examples of speech (and, indeed, writing) were the basis for the model. Secondly, such abstract approaches permit the theorist to ignore sound-based meaning-bearing elements of language, such as intonation, which are again less easy to formalise than text-based elements.

Much of a person’s identity and communicative force is carried by the vocal pattern that we associate with them, and many of the affective aspects of language reach the world via the slightest changes in voice quality. In teaching spoken language one might imagine these aspects would be seen as of highest importance. However, since most abstract language paradigms do not take into account or try to account for aspects of the dynamic, interpersonally oriented mode that is speech, the focus tends to fall on structural input, disengaged both from its discourse context and from its meaning-bearing ‘music’. In contrast to this, work that is ongoing in computer science and human–computer interaction is keen to better understand and incorporate findings about the links between communicative impact, affect, and prosody (for instance, Partala and Surakka (2004)). It will be interesting with the growth of multi-modal corpora and new techniques for searching these how far the findings of computer science, corpus linguistics, and the language classroom can be combined to provide insights that are eventually applicable to the spoken language curriculum.

The development of functional magnetic resonance imaging ‘fMRI’ technology linguistic research developed in the early years of this century has had an interesting effect on the study of spoken language. The capacity to link brain function to particular spoken stimuli has meant researchers can now build hypotheses to investigate questions about links between oral/aural input and events in the brain. The reason that this is a step change in the field is that earlier neurolinguistic work depended on making links between spoken events in the outside world and possible brain activity. This was often done by contrasting brain-damaged and non-brain-damaged speech performance. While this approach remains valid, the capacity to map and link spoken events and normal brain activity is an exciting new development for linguistics.
Concept 2.5  Functional magnetic resonance imaging (fMRI) and language

fMRI is a method by which activity in different parts of the brain can be shown as an image. The process is non-invasive as it scans from outside the body and analyses differences in blood flow. Changes in blood flow cause measurable fluctuations in oxygen levels and in turn its magnetic properties. The scanner translates these into data that are mapped on to particular areas in the brain. The assumption is that blood flow and neural activity happen hand in hand and therefore these images represent the physical location of the brain’s response to particular stimuli.

Work in this field on language developed rapidly from around 2000 and the process has been used for a wide variety of studies ranging from vocabulary (Ellis et al., 2006) to emotional responses to language (Beaucousin et al., 2007).

2.5 Levels of analysis

One of the difficulties in researching speech is the fact that, unlike written texts, the notion of a freestanding genre or clearly delimited sample to be investigated does not readily lend itself to speech. Whereas the researcher into writing can start, if they wish, from a relatively well-defined set of texts that clearly fit into a category (newspaper language, popular fiction, advertising texts, academic writing and so on), the researcher into speech will generally find no such helpful categories to hand. Writing presents itself in front of the researcher through the materiality of its visual medium. The researcher into speech must usually look beyond the discourse to the context in order to delimit the data under investigation and to ensure they are, for instance, comparing like with like.

The issue can be best understood by looking at a stretch of talk, and thinking about the different levels and perspectives through which it could be investigated.

Figure 2.1 shows a brief extract from interviews conducted to create a corpus of Singaporean English. The corpus was created primarily for the purposes of research into prosodic features. It presents a readily accessible set of transcripts alongside digital audio files of the original speech data. This leads to a preliminary, overarching comment that the transcript and the spoken data are not the same thing and should not be conflated. Researchers into speaking all need to reach a carefully thought out position in relation to the visible recording of their data in the written form as this is rarely a neutral process. This can be understood in terms of a metaphor of degrees of magnification. The sample provided captures a number of aspects of talk: socio-pragmatic relationships (interviewer/subject, lecturer/student),
structural features (turns, questions/answers, clauses) and acoustic data relating to the stream of speech (temporal information in seconds, onset of overlapping talk). However, these are only a small subset of features that may interest the researcher. At a level of higher magnification, someone might wish to represent relative loudness or pitch movement in individual words. The greater the acoustic information being captured, the greater the efforts involved in transcription. There is therefore always a relationship between the ‘magnification’ (level of detail captured in a transcription of speech) and the research focus. A researcher may make a very simple initial transcript showing no data such as overlapping talk or pauses if they are primarily interested in finding instances of a particular type of interaction (jokes, for example) and then increase the level of detail for those extracts. Taken seriously transcription is a powerful research tool and can reflect the perspectives and needs of the researcher. As it is never an entirely neutral process, it is good practice for the individual transcriber to cross-refer with other researchers when difficulties of interpretation arise or when new categories of talk are being investigated. O’Connell and Kowal (2009) provide a thoughtful summary of the development of transcription systems and issues to consider. The most commonly used system is often referred to as the ‘Jefferson method’ after the linguist Gail Jefferson who developed this. A definitive overview can be seen in Jefferson (2004). Setting on one side this methodological preliminary issue, the extract shown in Figure 2.1 could be the object of study at many levels and the following sections deal with each of these.

2.5.1 Analysing speaking skills at the level of discourse and social interaction

Discourse-level studies are interested in questions of how speakers interact with one another (for example, how they know when it is their turn to speak), and how talk is organised in particular kinds of patterns over long stretches of language (for example, how speakers structure their talk for
listeners so that they can follow changes in topic easily). At a wider level, researchers are often also interested in how, through talk, social features are expressed, such as identity, shared knowledge, or power relations. In the extract in Figure 2.1, the speakers are a lecturer (male) and a student (female). Their interaction takes place in a semi-formal interview setting. Their relationships, gender, and the interview context influence how they behave to one another conversationally. For instance, it is more likely that the lecturer/interviewer will initiate talk in this setting and it is likely that more of his discourse will be in the form of questions. Many disciplines outside linguistics are becoming increasingly interested in discourse analysis because of the insights it can give about how participants in a spoken interaction behave. For instance, researchers in the medical sciences may be interested in how to understand patient and practitioner relationships better in order to enhance training in communication for professionals and therefore the efficacy of treatment (see Salter et al. (2007) for an example of this work). Similarly, a wide-ranging recent summary of applications of linguistic analysis in the realm of business studies can be found in Bargiela-Chiappini et al. (2007).

During the 1970s and 1980s the main concern in the field was to consider where the discourse level of language fitted in with current views of language, and to what extent regularities or even ‘rules’ of interaction could be uncovered. This focus on rule-based paradigms reflected the dominant model for language that had grown up in the USA. Seminal work was carried out in America by conversational analysts who developed highly sophisticated systems for representing language features which had previously been studied very little, for example laughter or pauses or apparently trivial utterances, such as ‘uh huh’ or ‘oh’ (e.g. Schegloff, 1981). This detailed investigation into the mechanics of conversation led to concepts such as ‘openings’, ‘closings’, ‘pair parts’, ‘formulaic exchanges’ or the ‘transition-relevance point’ (TRP).

**Concept 2.6 Transition-relevance point (TRP)**

This is a moment in speaking when several linguistic features combine to signal to an interlocutor that they could take over the speaker role. In Anglophone cultures these tend to be the ends of clauses and are signalled by pitch, intonation, pace, micro-pausing as well as extra-linguistic features such as gaze. Next time you are in a free-flowing conversation you might like to stand back (or better still record a conversation) and see how speakers know that they can begin to speak without seeming to interrupt one another. For many learners of a language, ability to speak is not the factor which isolates them in a conversation. Rather it is the inability to ‘read’ the moments when they might be able to begin to speak.
In the UK key features of the structuring of discourse were investigated and notions such as ‘discourse markers’, ‘transactions’ and ‘exchanges’ were developed.

Concept 2.7 Discourse markers

These are words ‘outside’ clauses which carry little or no meaning in their own right but signal something to the listener about the structure or organisation of the talk, for example ‘right’ or ‘ok’ in English. As well as logical relations, discourse markers can signal more subtle aspects of talk: ‘well’ can indicate reservation or hesitation; ‘now’ can indicate a change in topic; ‘actually’ can mean many things including difference of opinion or correction or even defensiveness (cf. English cooking is very good. English cooking is very good, actually.). Because learners are usually taught a form of the language which is strongly influenced by written mode, spoken discourse markers are not given high prominence in a syllabus, if they are taught explicitly at all. This can leave a learner floundering both in terms of listening to conversation and taking part.

Both discourse analysis and conversation analysis have links to sociolinguistics in that they prefer not to deal with samples of language in isolation, and conversation analysis in particular is interested in the relations between interlocutors. Discourse analysis, however, has traditionally tended to concentrate on longer sections of language and focused on interrelations between different sections of text. Within this, the discourse analyst is interested in how speakers carry out functions of language and the choices made by them in different contexts.

In terms of the application of some of the main ideas of conversation and discourse analysis, but with a stronger focus on the former, Brown and Yule (1983) Teaching the Spoken Language: an approach based on the analysis of conversational English provided something of a bridge between the schools of thought outlined above and more practically classroom-oriented applications. Interestingly, despite the crucial aspects of speech that discourse-level studies have uncovered they have, overall, been very slow to trickle down into classroom teaching and published teaching materials in general. There have been books for teachers on the topic (for example, McCarthy, 1991, Discourse Analysis for Language Teachers or Evelyn Hatch’s, 1992, very different Discourse and Language Education). Carter, Hughes and McCarthy (2000) attempts to bring some of the complexities of spoken grammar in discourse to the classroom via grammar materials. Chapter 3 deals with this more fully.

Discourse analysis in the UK does have strong incidental links to the classroom, however, in that much of the most influential early work (for
example, Sinclair and Coulthard, 1975) was carried out on classroom interaction. These classic studies, which generated some of the fundamental categories of discourse analysis, were based on teacher–pupil talk.

During the 1990s and beyond there was increasing interest in the telecommunications and computing world that discourse analysis would solve problems of automation of human–computer understanding. This area has not achieved the early promise – humans are still constrained to limited lexical choices and clear talk in these contexts rather than the system being able to adjust to spontaneous talk. Nevertheless, it will be interesting to see what twenty-first-century discourse and conversation analysis can offer other disciplines and users wanting to apply the insights of linguistics to real-world applications.

The relationship between psychology and speech behaviour is another thread to research into global aspects of speech, and one which again links in with the bigger questions of how spoken language data relate to underlying linguistic systems, whether neurological, biological or genetic. Whereas discourse or conversation analysts will describe patterns of speech behaviour in order to uncover regularities in the organisation of spoken discourse, and will see these patterns as of interest in themselves, the psychologist will generally regard utterances as a source of evidence of mental or behavioural processes. So, for example, whereas a discourse or conversation analyst may look at a feature such as patterns of repetition in speech and see how far they can generalise about lexical repetition in its own right, the psychologist would investigate how such repetitions relate to how humans process complex utterances, or the timing and levels of pre-planning. For example, Clark and Wasow (1998) investigated a typical pattern of repetition, either PRONOUN + PAUSE/FILLER + PRONOUN (I uh I (think) . . .) or ARTICLE + PAUSE/FILLER + ARTICLE (The uh the (problem) . . .) and suggested that the different stages in this pattern related to the way in which speakers committed themselves to an utterance. They proposed that these items were an integral part of the underlying psychological processes by which utterances reach the world.

In terms of teaching languages, the fields of higher-level studies into speech described here open up several questions and ways forward, particularly in relation to uncovering differences between cultures in terms of how conversation is organised. This in turn can help learners and teachers understand potential pitfalls in language interaction that are not due to any grammatical mistake but different pragmatic and cultural expectations.

2.5.2 The research space at the level of language choices: grammar and vocabulary

A prominent strand of work on grammar and lexis that takes into account spoken mode has been developed through corpus studies. Douglas Biber's
work, and most notably his influential *Variation Across Speech and Writing* (Biber, 1988), gives a strongly data-oriented analysis of a wide variety of spoken and written sources, concluding that certain grammatical features cluster together to make up the distinctive style of a spoken or written genre. These features in turn map on to dimensions of contrast, such as whether the language is concerned with conveying information or is more inter-personally oriented. Rather than suggesting a simple binary division between speech and writing, Biber suggested that there were patterns of probability among language features that show statistical regularity in how they co-occur in spoken and written genres.

There have been two major strands of work developed from this approach: applications of the register analysis in discrete fields and genres, and more theoretical and detailed insights about general language features. In terms of the former, this has been taken forward in, for instance, Douglas Biber's own work on English for special purposes including university writing and speaking (Biber, 2006), and with colleagues on English language assessment (Biber *et al.*, 2004), and the large body of work around the Michigan Corpus of Spoken Academic English (‘MICASE’). In terms of the latter there has been, for instance, some work on language change as a feature capable of being investigated through corpora (see Concept 2.8 describing the process known as ‘grammaticalisation’).

---

**Concept 2.8 Grammaticalisation**

This is a concept associated with the study of language change. It is based on an underlying idea that words in a language can be categorised as primarily carrying semantic meaning or as primarily carrying out syntactic functions. The verb ‘walk’ would be in the first category and the verb ‘have’ would be in both categories. It has a lexical meaning when used in the sentence ‘I have a brother’ and has less semantic load when used in its auxiliary functions, as in ‘I have broken my arm’. At different phases in the evolution of a language, words can change from lexical to grammatical functions and this process is called grammaticalisation. An example that is often given is the expansion of the lexical word ‘back’ as part of the body into an adverbial to indicate past time as in ‘Back then . . .’. Lindquist and Mair (2004) provide a collection of research papers on what is also termed ‘historical corpus linguistics’.

---

Early corpus-based work helped to inform research into speech on detailed aspects such as tenses (Aarts and Meyer, 1995), vocabulary (Stenström, 1990), clauses (Nelson, 1997), and ellipsis (Meyer, 1995). More recently, work has involved broader questions of the extent of generalisability and
applicability of findings (for example, Conrad (2000), Mindt (2002), Baldwin et al. (2005)). Academics also began working on bridging the gap between corpus findings and the classroom (e.g. Johns, 1991; Tribble and Jones, 1990; Knowles, 1990; and most relevantly for the spoken mode, Svartvik, 1991) as soon as large corpora of speech became readily available. More recently scholars have focused attention on the interrelations between second language acquisition theories, or approaches to teaching, and the grammatical frequencies found in corpora (for example, Biber and Reppen (2002), Anderson (2007), Barbieri and Eckhardt (2007)). However, while these can provide some interesting ideas for classroom activities, there is still a gap between findings on the realities of grammatical and lexical choices in spontaneous talk, and what is presented in published material.

On the other hand, for the teacher or researcher starting out in research into spoken mode, gaining access to spoken corpora has become relatively straightforward. These resources range from sampling a larger corpus on-line (for example, the British National Corpus at: http://sara.natcorp.ox.ac.uk/lookup.html) to ordering a corpus or a sample on a CD-ROM (for example via ICAME at http://www.hd.uib.no/icame/newcd.htm) to gaining full access for scholarly work either via the corpus designers or publishers associated with the project. For further details, see Chapter 9.

**Quote 2.4** The linearity of speech and how grammar is conceptualised

One fundamental difference between spoken and written language has to do with the ‘linearity’ of speaking in time, in that the temporal structure of speaking is inherently the outcome of an interactive process between speaker and listener. But despite the status of ‘linearity’ as one of Saussure’s fundamental principles, in practice little more than lip-service is paid to the temporality of spoken language, which is treated as having few if any consequences for syntactic analysis. It is trivial to point out that a structuralist definition of the sentence is incompatible with an on-line model of syntax processing. A structuralist analysis, even of ostensibly spoken language, is carried out not from a real-time emergence perspective but as if it were – like a written text – a finished product.

(Auer, 2009)

Another area of work on spoken grammar looks at the interface between prosody and syntax. In the latter decades of the twentieth century, there were some notable attempts to set aside descriptive and prescriptive grammars and to incorporate prosodic elements into the analysis of the form. David Brazil’s (1995) *A Grammar of Speech* was particularly unusual because the linear nature of speech production is taken seriously and talk is described in terms of having ‘purpose’. It is interesting to note that, although Brazil’s
book is highly regarded, the teaching profession as a whole has found it difficult to assimilate many of the principles that underpin the work. The class text on pronunciation for advanced learners of English also by David Brazil (Brazil, 1994) presents a similar unification of discourse-level and other meaning-bearing language features. Hence, in many ways Brazil's work represents one of the most consistent attempts to look at the spoken form on its own terms.

Klein and Purdue (1992) *Utterance Structure: developing grammars again* was another exception to the rule. In this deliberately provocative and stimulating work the authors take issue with many of the assumptions of second language acquisition and base their analysis on the notion that learner utterances are a language to be studied in their own right (rather than in relation to a ‘target’ language). The book is of particular interest to the researcher into speech because it takes a strongly empirical (data-oriented) stand and builds the discussion on real utterances. However, the researcher new to speech research studies, reading either of these books, should realise that while both of them are fascinating neither of them have succeeded in becoming centrally accepted into the inner circle of applied linguistics, partly due to their (intentional) lack of overlap with the theories underlying the field more generally.

During the last decade of the twentieth century and first decade of the following one, a growing trend in the research into spoken grammar has been the field known as interactional linguistics. This is a highly relevant area for the researcher interested in approaching spoken data in terms of the dynamic and ‘real time’ aspects that tend to be lost when the spoken form is examined entirely through the lens of the written mode (see also Quote 2.4).

Concept 2.9 **Interactional linguistics**

This is a branch of linguistics that is closely related to conversation analysis. Whereas CA analyses speech data to better understand patterns of talk and the social aspects of talk that are revealed by interactions between participants, IL uses many of the same methods of analysis but the emphasis is on insights about language itself that can be gained from examining talk in interaction.

Work in IL has a strong connection with the Santa Barbara Corpus of Spoken English and the development of this by John Du Bois and colleagues was something of a catalyst for work at the interface between discourse and grammar. A bibliography of work in the area can be found at http://www.linguistics.ucsb.edu/faculty/sathomps/bibliographies/bibliog-interactional-linguistics.htm and the approach has been used to allow better understanding of grammatical variation that appears sensitive to interactional context (for example, how talk is organised around repetition (Bybee, 2006), syntactic constraints on offers of assistance (Curl, 2006) and so on).
Interactional linguistics is closely related to conversation analysis and approaches the emergent nature of grammatical elements from that perspective. There is also work on the interface between prosody and syntax that is aligned towards the experimental paradigm, for example Grosjean (1983) and Grosjean and Hirt (1996) which looked at the issue of how listeners predict the end of a clause from acoustic information earlier in the clause; or Marsi et al. (2002) who carried out work on this area in text-to-speech recognition.

2.5.3 The research space at the level of speech production: fluency and pronunciation

Traditionally, in the mid to late twentieth century, a great deal of research into phonology was undertaken to find evidence of an underlying system. This was generally carried out in relation to the model of language that tended to dominate at that time, and to an extent still does: the transformational or the universal grammar paradigm (see, for example, Nestor and Vogel (2007), Prosodic Phonology). Within this paradigm, some consideration was given to the interplay between the different levels, and the direction of influence between them. However, the aim of the work was generally to find evidence of internal language knowledge rather than to describe the system for any applied purpose. For example, Berg and Hassan (1996) examined speaker errors in three languages but were less interested in classifying or explaining the errors than in gaining insight into the ‘mapping’ or hierarchical conceptualisation of speakers’ linguistic knowledge in the three cases.

In contrast to this approach, work into pedagogically related phonology has been carried out within a much less theoretically oriented, and generally an experimentally based, framework. Experimental-style research into the teaching of pronunciation is an area that can provide clearly relevant results for the classroom and Derwing and Monro (2005) provide a good summary of the research and teaching interface in this area. Keeping up to date with these kinds of results can help teachers plan the balance of the speaking syllabus, and can also account for contrasts between student progress in and outside the classroom, for example, if a student appears to be making good progress in pronunciation in controlled circumstances but remains difficult to understand when producing longer sections of speech.

In contrast, and at a more theoretical level, an emergent area in research of prosody has been studies of English spoken as a lingua franca (Concept 2.10). This work emerged during the early years of this century and is an area that has provided lively debate about standards, ownership of the language, and the balance between intelligibility versus a native speaker model in teaching speaking. Work such as Jenkins (2000) crystallised some of these ideas and a readable summary of the issues can be found in Pickering (2006). Chapter 6 returns to this topic in detail.
Concept 2.10  **Lingua franca, intelligibility, and common core features**

A lingua franca is a form of language used as a common one between speakers of different languages. It is, by its nature, associated with particular domains where people need to use language to talk to people they are unfamiliar with and to carry out functions beyond the family and other local domains. Latin was a dominant lingua franca in mediaeval Europe where it was used for religious and scholarly purposes alongside local ‘mother tongue’ languages. English is currently often used as a common language for business, academic and other purposes and a branch of applied linguistics research is on ‘ELF’ or English as a lingua franca (Seidlhofer, 2001). This work relates to discussions surrounding the importance of intelligibility over accuracy (as defined by Anglophone norms), and what common core of grammar and vocabulary is required for users to carry out communicative functions in what is their second or third language.

**Summary**

This chapter reviewed some classical approaches to the research process and addressed the particular problems for the researcher working with spoken forms of language. Beside the issue of the lack of extensive work on the spoken form in its own right, I raised the question of the role of speech data in language theory generally, and of the attitude to situated spoken discourse as the basis for generalisations about language. In relation to research based around teaching the spoken form, the further issue of the cultural and pragmatic problems raised by real speech data was aired.

**Further reading**


