

Part I
Linguistics-Applied (L-A)

Introduction to Part I: Linguistics Applied (L-A)

ALAN DAVIES

We have argued in our general introduction that while the distinction between Linguistics Applied (L-A) and Applied Linguistics (A-L) is fugitive, it remains necessary and that it is at its most obvious in the orientation of the researchers, why they are investigating a problem and collecting their data. If they regard themselves as linguists applying linguistics because they wish to validate a theory, that is linguistics applied (L-A). If they see themselves as applied linguists because they seek a practical answer to a language problem, that is applied linguistics (A-L). Having made that distinction, we offered the caveat: "We do, of course, recognize that in some, perhaps many, cases the researcher will have both interests at heart." We should also point out that the orientation of the researchers, how they regard themselves, what it is they wish to achieve, is not always obvious. Even when asked, researchers may not be clear.

The L-A chapters that follow in Part I present a tendency, a tendency toward the investigation of language using linguistic or other modes of investigation. What I propose to do is to group the 16 chapters in Part I into six sections; the sections themselves providing a cline from closest to the linguistics of language to the more distant connection. Thus in Section 1 we have the Liddicoat and Curnow chapter (on descriptive linguistics) which offers a descriptive apparatus for the linguistic areas of grammar and phonology. Such a chapter could with ease fit into a handbook dealing with linguistic descriptions. No problem there! The border between L-A and A-L is not marked and just as A-L needs linguistics, so too L-A requires a means of handling its application. Also in Section 1 is the Kirkness chapter on lexicography. The purpose of the Liddicoat and Curnow chapter is "to introduce applied linguists to the broad themes and general concepts with which linguists work in developing descriptive accounts of language". Applied linguists, they argue, need "a certain level of familiarity with the principles of linguistics" so that "the work of applied linguistics can be carried out in an informed and principled way" For Liddicoat and Curnow linguistics is system and while this may not be the driving force in applied linguistics, applied linguists must

come to grips with language as a system since “linguistic and language description is basic to applied linguists’ work”. To that end, Liddicoat and Curnow provide an introduction to phonetics/phonology, grammar, and semantics.

Their chapter therefore is linguistics for applied linguists and as such very much at the linguistic end of L-A. In his chapter on lexicography, Alan Kirkness is similarly more linguistic than applied. Even so, as he points out, lexicology operates at the level of particular languages and while, in doing so, it makes use of linguistic procedures and constructs, it is powerfully concerned with the uses made of lexical research. Kirkness maintains that there is and always has been at the heart of lexicology an interest in application. Most particularly in dictionary making for various purposes. And he ends with a compelling plea for a close link between lexicology and lexicography, between the theoretical and the practical, between the linguistic and the applied. What that means is that lexicology belongs, in our terms, to L-A and, within L-A stands at the linguistic end of that approach.

Section 2 consists of chapters that investigate language in terms of the uses that are made of it. For David Birdsong, second language acquisition (SLA, or, as he puts it, L2A) is “a central concern of Applied Linguistics (or more precisely . . . of Linguistics Applied)”. Such a view is orthodox among SLA researchers: for them (as for Birdsong), the purpose of SLA research is to further our linguistic understanding, not to develop more effective ways of learning and teaching languages. Of course, such spin-offs may follow, but they would be incidental to the role they envision for SLA research, to model and promote our understanding of language and its acquisition. Birdsong’s take on the topic is not mainstream in that instead of the more usual account of initial SLA, he discusses “the end state” or “ultimate attainment.” He reminds us that “ultimate attainment data are invaluable for ongoing mainstream research in L2A theory, in that they afford unique perspectives on the limits of L2A . . . Clearly, for educators and social-policy makers, as well as for theorists, it is of compelling interest to know more about the rate of native-like attainment”. Such an approach could illumine “the most basic issue on L2A research . . . whether the difference in ends (i.e. final states) implies different means (i.e. learning procedures). As well as the L2-L1 comparison, Birdsong addresses the age factor in SLA. While his orientation is very obviously L-A, it is all too clear that his interest in the basic issues of L1-L2 and of age of acquisition are also of central interest to A-L.

For Mike Stubbs, the advent of computerized corpora provides a kind of paradigm shift in linguistic description and in our understanding of language and its development over time. What corpus study does is to bring together as parameters (and therefore unfalsifiable) populations of language tokens across individuals. In other words, what linguistics has always done manually and partially. For our purposes, then, corpus study necessarily falls into the L-A area: “no linguist” Stubbs claims “can now ignore corpus data”. But does corpus study do more, does it have any applied reach? For Stubbs there are areas of application: he mentions language teaching, lexicography, translation

studies, stylistics, forensic linguistics, cultural representation, and psycholinguistics. But his claims are modest. While he is unapologetic regarding the value of corpus study for linguistic descriptions (he calls himself an enthusiast here), he offers a conservative view of applications “arguing that applications are indirect, and that before findings can be applied to real-world problems, they require careful interpretation”. What we can be sure of is that corpus studies, like lexicography, like discourse analysis, are good for linguistics. Are they good for applied linguistics?

Trappes-Lomax reminds us that discourse analysis is practiced by scholars in many disciplines and not only by those working in linguistics and applied linguistics. The “linguistic turn” in the social sciences has largely been about this continuing interest in discourse analysis, which recognizes the value of non-experimental and non-quantitative methods in managing evidence. Trappes-Lomax takes us through the five areas he terms “focal issues” in discourse analysis: these are interaction, context, function, instrumentalities, and text. He defines discourse analysis as “the study of language viewed communicatively and/or of communication viewed linguistically”. Such a wide lens may be too generous since it can be seen as inflating the claim to our attention of discourse analysis by equating it with applied linguistics. There is a warning here. As with SLA (and indeed critical applied linguistics), the excitement and enthusiasm for the research interest may encourage inflation in the value of the research such that then applied linguistics becomes wholly SLA or CAL, or, in this case discourse analysis. But what cannot be denied is Trappes-Lomax’s claim that discourse analysis is necessary “to our understanding of language, of society, and of ourselves as human beings . . . it is useful – in an ever expanding range of practical and socially beneficial activities . . . (and) it is . . . endlessly interesting”.

One of the ways in which linguistic theory can be applied to language problems is by differing ways of linguistic description: we saw that in Section 1, particularly with the Liddicoat and Curnow chapter which provides a methodology for description at a level more abstract than an individual language. Thus the writing of a grammar of English (or of Japanese) would be a way of describing language at a somewhat less abstract level. The chapter by Sutton-Spence and Woll therefore belongs here since it concerns the description of a particular language, in this case British Sign Language (BSL), and what the chapter discusses is how linguistic procedures and methods can be implemented in order to establish a description of BSL. For Sutton-Spence and Woll, BSL is a minority language; but so of course are many oral languages. It is British (as is English, as are the Celtic languages . . .), it has its own speech community, again like all oral languages, but uniquely it is a visual language. In other words, for Sutton-Spence and Woll, BSL is fundamentally a language: the fact that it uses visemes rather than phonemes is, in a profound sense, trivial.

In Section 3 we examine approaches that uncover the connections between speakers and their language, thus Giles and Billings, Schmid and de Bot,

Kramsch, and Gardner. In their chapter *ASSESSING LANGUAGE ATTITUDES: SPEAKER EVALUATION STUDIES*, Giles and Billings explore the interaction between language, communications, and social judgments, recognizing, as they do, that “the effects of language on social judgment is an integral part of uncovering the communication process”. What speakers use language for, in other words, is to make judgments about their interlocutors: the fact that social judgments are often stereotypical emphasizes that it is a language rather than a linguistic evaluation that is being made.

Schmid and de Bot examine in their chapter various approaches to the study of language attrition, pointing out that just as languages are gained/acquired so they are lost: they investigate the phenomenon of loss at the individual and the community level, noting that languages are lost both deliberately and non-deliberately, through migration, contact, aging, and trauma. How far language attrition and SLA are mirror images remains an intriguing question. For our purposes, what Schmid and deBot (like Giles and Billings) are centrally concerned with is the ways in which speakers relate to their (and others’) language.

As well as viewing language as a resource and/or commodity, as Schmid and de Bot do, we can also regard it as both vehicle and simulation of thought and culture. This is the concern of Claire Kramsch in her chapter *LANGUAGE, THOUGHT, AND CULTURE*. Kramsch traces the progress of applied linguistics from its universalist certainty in the 1950s and 1960s through to its more questioning, context-sensitive relativism of today. She takes three areas to demonstrate this shift in linguistics: semantic relativism, linguistic relativism, and discursive relativism and then maintains that this shift has followed on, lagging behind, in applied linguistics. This has, she maintains, affected the orientation of speakers to their language above all in language education: “language relativity suggests reorienting the focus of language teachers from what they do to who they are”.

Gardner’s chapter on conversation analysis (CA) provides another take on the ways in which speakers use language: as we have seen, they form attitudes toward it, they view it as part culture and part culture bearing, and they lose it. In all cases, what the analyst is doing is focusing on the interaction between the speaker and the language. Here too in Gardner’s account of conversation analysis we see a similar focusing. Gardner shows how CA borrowed three basic themes from ethnomethodology: accountability, reflexivity, and indexicality. As well as being grammatical and appropriate, speakers are accountable, reflexive, and indexical for the purpose of effective interaction. And it is these themes that CA studies, what Gardner refers to as “the complexities, local design and quiddity of instances of talk”, in other words, how language is used to create language meanings. To what extent the systematic use of conversation should take account of “local design and quiddity of instances” remains unclear. Gardner appears not to take the Kramsch view, and concludes that “ordinary conversation is likely, at least in many of its instances, to be universal”.

In Section 4 we place three chapters that concern various functional uses of language: LANGUAGE AND THE LAW, LANGUAGE AND GENDER, and STYLISTICS. In all three cases, while the traffic is both ways, what seems primary is the light thrown by these functions on the language itself. What distinguishes Section 4 from Section 3 is that while Section 3 deals with applied linguistics in terms of language, Section 4 concerns applied linguistics in terms of language use. Gibbons, writing on language and the law (also termed “forensic linguistics”) proposes that the law is an applied linguistic issue because the law (unlike, say, medicine) is based on and mediated through language. His chapter examines four sources of the problems that arise: the “genre” issue (“the specialized text structure and procedures used in the law”), the “writteness” of legal documents (that is, that they are accessible only through reading), the “technicality” of legal discourse (rendering its understanding inaccessible to non-lawyers), and the “interpersonal arena” (given the power imbalance in legal processes). Gibbons presents legal language as a type of code: making that code accessible to those in need (“people who cannot understand the legislation impacting on their lives, witnesses whose testimony is distorted by linguistic pressure tactics, minorities whose language cannot be used or who are subjected to group vilification, or the guilty or innocent convicted by language evidence” is a proper task for L-A.

Susan Ehrlich (LANGUAGE AND GENDER) maintains that people *do* gender through the linguistic choices they make. Gendered language is therefore a (deliberate) choice made by speakers. In the same way that lawyers construct their legal identity through language, so do men and women construct their (gendered) identity through linguistic practices. Interestingly, Ehrlich makes a convincing case for bringing together the two main areas of language and gender research: the study of language use and the study of sexist language. Her argument is that the one is the product of the other, that sexism is an act (doing things with words) with outcomes affecting identity and judgments. This is a relativist neo-Whorfian view and fits well with the Kramsch discussion above on language, thought, and culture.

McRae and Clark recognize that stylistics “has proved notoriously difficult to define, since it functions as an umbrella term”. For our purposes, what is of interest in stylistics is its concern with a particular language use (its textness, originally entirely literary, more recently quite general). We might think that stylistics would make a more powerful impact if it was still wholly concerned with literary texts. Even more than language and the law, stylistics is language bound. The authors explain how valuable stylistics can be in the teaching of literature as a foreign language, hardly surprising given the long centuries during which literature featured as a main (perhaps the main) component of language teaching. What this chapter does is to make a case for stylistics as a way of applying linguistics to the educational study and understanding of (literary) texts.

Section 5 contains two chapters dealing with the influence of language in external affairs, notably in politics. Thus Joseph proposes that “the study of

language and politics is aimed at understanding the role of linguistic communication in the functioning of social units, and how this role shapes language itself". Language influences the political; equally the political influences language. In the case of language and the law, language is the medium of the law; here, in the case of language and politics, language is substance as well as medium. One of the examples Joseph quotes is that of the globalization of English as an instrument of linguistic imperialism. It is this topic of the spread of English (here called world Englishes) that Kingsley Bolton addresses in his chapter. Bolton helpfully points to the dilemma of applied linguistics in approaching the fact of World Englishes where "considerable problems for applied linguistics still exist in the area of pedagogical principles and practices. He refers to local attitudes and official practices, noting that the way ahead may require "new and creative approaches" which might mean the reorientation of the whole concept of World Englishes from its current L-A status to one that is more A-L.

The last section (Section 6) in this part of the volume has only one chapter. Kanavillil Rajagopalan's *THE PHILOSOPHY OF APPLIED LINGUISTICS*. His chapter exercises a Janus-like function in the volume, looking back at L-A and forward to A-L. This encompassing embrace is provided by Rajagopalan's historical overview, an account both of periods and ideas. The chapter charts the beginnings in the mid twentieth century when linguistics was the driving force in applied linguistics, through the Chomskyan revolution (bringing with it the long-term emphasis on SLA research and what Rajagopalan calls "the apotheosis of the native speaker"). And so to the sunny uplands of interdisciplinarity where L-A yields to A-L, the underlying topic of our Part II. The story does not end on those uplands, as Rajagopalan makes clear, but moves on to a putative post-A-L, which is what critical applied linguistics claims to be. It is not accidental therefore that the last chapter in our Part II (A-L) and in the volume deals with critical applied linguistics. That is for later. We turn now to the 16 chapters in Part I, the L-A approach to applied linguistics.

1 Language Descriptions

ANTHONY J. LIDDICOAT AND
TIMOTHY J. CURNOW

1.1 Introduction

The importance of language description in applied linguistics has sometimes been questioned (e.g. by Widdowson, 1979, 1980) because of a perception that the theoretical insights of descriptive linguistics are different from the practical needs of language pedagogy. Linguistics has increasingly separated itself from a prescriptive view of language, which formulates rules for what should be said or written, in favor of a descriptive view, which seeks to record the language which people actually use. Contemporary language description, therefore, takes a synchronic approach, that is, language is described as it is at a particular moment in time and does not incorporate the history of the language (diachrony), although languages do of course change over time.

The descriptive view has led linguists to new insights about language and new ways of talking about and defining units of language. However, in many cases applied linguistics has required a prescriptive grammar recognizing that language teaching is frequently a case of teaching what should be done (Odlin, 1994). In other words, pedagogical grammar has been equated with prescriptive grammar. Pedagogical grammars have tended to adhere to the concepts and terminology of traditional grammar, based on the linguistic categories found in Latin and Ancient Greek, and, especially in the case of first language teaching, often have had a diachronic perspective, favoring rules based on earlier forms of the language. Recently, however, especially with the introduction of corpus-based materials into language classrooms, pedagogical grammar has taken on a more descriptive focus, with learners being required to deduce rules from linguistic data (cf. Tomlin, 1994; Kennedy & Miceli, 2001).

At the same time, applied linguistics itself is not entirely a pedagogy-focused discipline and many areas of applied linguistics have pursued language description as a central feature of their work. This is especially true of first and second language acquisition, where much work has been done on the description of learner grammars. Moreover, language standardization and vernacular

language literacy have both faced the challenges involved in bridging the divide between description and prescription and the development of pedagogical grammars from language descriptions.

Applied linguistics is focused on language, and while many applied linguists are not directly involved with language description, knowledge of the approaches and concepts of linguistic description is an important part of the working knowledge of any applied linguist (cf. Stubbs, 1986). In this chapter, we aim to give a brief overview of the main dimensions of linguistic description and the key concepts involved. The terms we use here are generally accepted, however particular theories may use different terms or define these terms in slightly different ways.

Descriptions of language are often divided into a number of categories and each of these categories has its own principles, concepts, and objects of study. For this paper we have separated language description into the study of the sounds of language (phonetics and phonology), language structures (morphology, syntax, and information structure), and meaning (semantics).

1.2 Phonetics

Most languages are transmitted by sounds and one of the most obvious differences between languages is that they sound different. The study of the sounds that human beings make in their languages is known as phonetics. While sign languages, such as British Sign Language and American Sign Language, are clearly not transmitted by sound, there are units in sign languages which correspond to phonetics and phonology, but these will not be discussed here (other areas of language description apply equally to spoken and sign languages).

1.2.1 *Transcribing sounds*

We are used to the idea of representing language in writing; however, conventional writing systems are not adequate to represent sounds. We need only consider the problems inherent in English spellings such as *cough*, *dough*, and *through* or the different pronunciations of words in US and UK English to see the problems involved in using conventional spellings to represent sounds: the sounds of a language are not the same as the letters of a language even in languages with much less irregularity than English. To overcome the deficiencies of conventional spellings, linguists use a phonetic alphabet such as the International Phonetic Alphabet (IPA) to represent sounds. IPA has over 100 symbols each representing different possible sounds. Phonetic transcriptions are usually written between square brackets.

In transcribing language we can use either a narrow transcription or a broad transcription. A narrow transcription contains as much information as possible and records very minor differences between sounds, while a broad transcription contains less information and records only some differences between

sounds. For example, a broad transcription of the word *pea* might capture the fact that it has two main sounds [pi], a narrower transcription might show that the consonant is actually unvoiced and aspirated and the vowel is long [p^hi:]. A very narrow transcription might include features of voice quality. Narrow transcriptions are very important in areas such as speech pathology or forensic phonetics where minor differences between sounds are important, but in most cases broad transcriptions are adequate for describing languages.

1.2.2 *The sounds of language*

The core of phonetics is to identify the characteristics of the sounds which human beings can use in language. Sounds can basically be divided into two types: vowels and consonants. Vowels are produced by altering the shape of the vocal tract by the positioning of the tongue and lips. Consonants are sounds which are produced by a partial or complete constriction of the vocal tract.

1.2.2.1 *Vowels*

Vowels are usually described by reference to five criteria, and these are adequate as a basic point of reference, although some vowel sounds require more specification:

- 1 the height reached by the highest point of the tongue (high, mid, low),
- 2 the part of the tongue which is raised (front, center, back),
- 3 the shape formed by the lips (unrounded or spread, rounded),
- 4 the position of the soft palate (raised for oral vowels, lowered for nasal vowels),
- 5 the duration of the vowel (short, long).

Using these features, linguists have constructed a set of standard reference points for describing vowels. These are called the cardinal vowels and are usually shown on a schematized representation of the mouth, as in Figure 1.1. In this diagram, the first vowel of each pair is rounded, the second unrounded, and all vowels are short. To show a long vowel, the symbol [:] is written after the vowel. The cardinal vowels are not all of the vowels found in human languages and some, such as [œ], are not even very common. There are many intermediate vowel sounds which fall between the cardinal vowel points, as we can see if we look at the vowel chart for English in Figure 1.2.

English vowels are usually oral. In French, there is a regular series of nasal vowels, that is, vowels which are produced by passing air through the nasal cavity by lowering the soft palate, shown by the symbol [̃] written over the vowel. The nasal vowels of French are [ã] *vent* 'wind', [ɛ̃] *pain* 'bread', [ɔ̃] *pont* 'bridge' and for some speakers [œ̃] *un* 'one'. Another feature of English is that front vowels are unrounded and back vowels are rounded, but this is not true of all languages. French, for example, has a series of front rounded vowels: [y] *tu* 'you', [ø] *peu* 'few' and [œ] *peur* 'fear'.

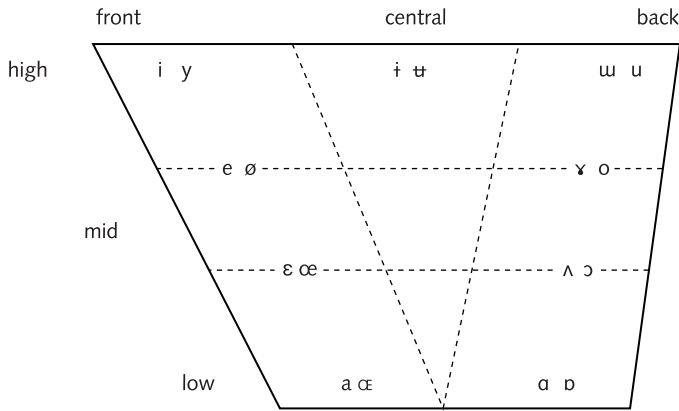
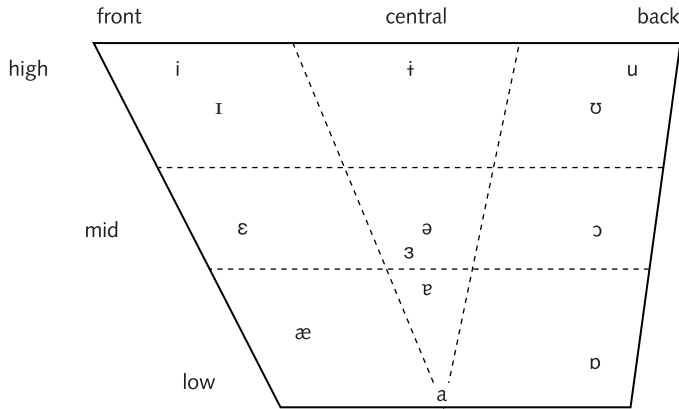


Figure 1.1 Cardinal vowels



Symbol	Example	Symbol	Example	Symbol	Example
i	<i>be<u>a</u>d</i> [bi:d]	ɨ	<i>piec<u>e</u>s</i> [pi:sɪz]	u	<i>fo<u>o</u>d</i> [fu:d]
ɪ	<i>b<u>i</u>d</i> [bɪd]	ə	<i>ab<u>o</u>ut</i> [əbaʊt]	ʊ	<i>put</i> [pʊt]
ε	<i>b<u>e</u>d</i> [bɛd]	ɜ	<i>w<u>e</u>re</i> [wɜ:]	ɔ	<i>po<u>r</u>t</i> [pɔ:t]
æ	<i>b<u>a</u>d</i> [bæd]	ɐ	<i>b<u>u</u>t</i> [bɐt]	ɒ	<i>po<u>t</u></i> [pɒt]
ɑ	<i>pa<u>r</u>t</i> [pa:t]				

Figure 1.2 English vowels (southern British variety)

In some languages vowels may be voiceless, that is, they are made without vibrating the vocal cords. This is shown by the symbol [◌̥] written under the vowel, as in Japanese *hito* 'person' [çito̥], *suki* 'like' [sʊki̥].

1.2.2.2 Diphthongs

Diphthongs are vowels in which the tongue starts in one position and moves to another. Diphthongs are very common in English:

tile [taɪl] *tail* [teɪl] *comb* [koum] *shout* [ʃaʊt]
toy [toi] *hair* [hɛə] *here* [hiə] *tour* [tʊə]

It is possible to have vowel sounds in which the tongue moves to more than one additional position during articulation. Some varieties of English in the UK, Australia, and New Zealand have triphthongs with three different tongue positions, for example:

fire [faɪə] *hour* [aʊə]

1.2.2.3 Consonants

Consonant sounds have three basic features in their articulation: place of articulation, manner of articulation, and voicing.

Place of articulation refers to where in the vocal tract the constriction is made using the tongue or other parts of the mouth. The most commonly used places of articulation are shown in Table 1.1. Manner of articulation refers to how the constriction is produced. The most common manners of articulation are shown in Table 1.2.

When air is passed through the larynx, the vocal cords may either be spread or drawn together. When the vocal cords are drawn together they create a vibration and sounds made with such a vibration are called voiced sounds (e.g. English *z*, *v*), while sounds made with spread vocal cords are called voiceless (e.g. English *s*, *f*). In reality the situation is a bit more complex than a simple distinction between voiced and voiceless consonants, especially in the

Table 1.1 Places of articulation for consonants

Place of articulation	Articulators	Examples
Bilabials	Both lips	English <i>p</i> , <i>b</i> , <i>m</i>
Labio-dental	Upper teeth and the lower lip	English <i>f</i> , <i>v</i>
Dental	Upper teeth and tongue	French <i>t</i> , <i>d</i>
Interdental	Tongue between the teeth	English <i>th</i>
Alveolar	Tongue and the alveolar ridge (the bony ridge just behind the upper teeth)	English <i>t</i> , <i>d</i>
Postalveolar	Tongue and the front edge of the hard palate	English <i>sh</i> , <i>r</i> in some varieties
Palatal	Tongue and the hard palate	Italian <i>gn</i> , <i>gl</i> , English <i>y</i>
Velar	Tongue and the soft palate	English <i>k</i> , <i>g</i> , <i>ng</i>
Uvular	Tongue and the uvula	French <i>r</i>
Pharyngeal	Pharynx wall	Arabic <i>ع</i>
Glottal	Glottis (vocal folds)	English <i>h</i> , Samoan'

Table 1.2 Manner of articulation for consonants

Manner of articulation	Type of constriction	Examples
Stop	Complete blockage of air flow	English <i>b, d, g</i>
Fricative	Turbulent airflow produced by forcing air through a narrow aperture	English <i>f, s</i>
Approximant	Partial constriction of airflow, but without turbulence	English <i>l, w, y</i>
Affricate	Blockage of airstream with a delayed release of the block creating turbulence	English <i>ch, j</i>
Nasal	Blocking of the oral cavity to force air through the nasal cavity	English <i>m, n, ŋ</i>
Lateral	Air flows around the sides of the tongue	English <i>l</i>
Trill	Repeated interruption of the airflow as the result of an articulator vibrating	Spanish <i>rr</i> , Italian <i>r</i>
Flap or tap	Very brief blockage of the airflow	Spanish <i>r</i> , Japanese <i>r</i>

case of stops. When a stop is produced, it is possible that voicing will occur throughout the articulation of the stop (voiced), at the moment that the blockage of the airflow is released (unvoiced) or after the moment of release (aspirated). This is known as voice onset time. In some languages such as Khmer, all three voicing contrasts are found: e.g., *baang* /ba:ŋ/ 'older sibling', *paang* /pa:ŋ/ 'to expect', *phaang* /p^ha:ŋ/ 'too'. English makes a distinction between aspirated and unaspirated stops only, while French distinguishes between voiced and unvoiced stops. The IPA symbols for the main consonants are given in Table 1.3. In addition, in some languages consonants may be long or short: e.g., Italian *notte* 'nights', *note* 'notes'. This is in IPA shown by reduplicating the consonant: [nɔtːe], [note].

1.2.2.4 Suprasegmentals

Individual sounds are considered to be discrete segments, however some of the sound properties of languages extend over more than one segment. These are known as suprasegmentals and include stress, pitch, and tone. Stress, tone, and pitch are assigned to syllables or even longer combinations of sounds rather than to individual sounds.

Stress refers to the prominence of a particular syllable in a word, usually the result of a difference in the loudness, pitch, and/or duration. For example, the

underlined syllables of the English words *develop* [dɪ'vɛləp], *language* ['læŋgwɪdʒ] and *about* [ə'baʊt] have greater prominence than the other syllables. These underlined syllables are stressed (shown with ['] before the syllable in IPA transcription) and the less prominent ones are unstressed. In English, unstressed syllables are often reduced, as in *about*, where the unstressed vowel is pronounced as [ə]. Longer words may have a secondary stress, a syllable with more prominence than an unstressed syllable, but less prominence than a stressed syllable, as in the underlined syllables of *controversial* [ˌkɒnt'rɪvɜːʃl] and *misdeemeanour* [ˌmɪsdə'miːnə]. Secondary stress is marked by [,] before the syllable.

Tone is a particular pitch which is assigned to the articulation of a syllable. In tone languages such as Mandarin Chinese these changes of pitch serve to distinguish individual words. In Mandarin there are four different tones:

high level	mā	'mother'
rising	má	'hemp'
falling	mà	'scold'
fall-rise	mǎ	'horse'

Some languages have a larger number of tones. For example, Thai has five tones and Cantonese has nine tones.

In some languages, known as pitch accent languages, pitch works in a slightly different way. In these languages, there are commonly two pitches – high (H) and low (L) – either of which is assigned to an individual syllable. In polysyllabic words, the pitch may vary across the word. This can be seen in the following Japanese words:

HL	<i>kaki</i>	'oyster'
LH	<i>kaki</i>	'fence'

Stress and pitch may also be assigned to larger units of language, such as sentences, in which case we talk about sentence stress and intonation (Cruttenden, 1997). English uses both of these. Sentence stress involves giving additional prominence to a particular lexical item in the sentence. For example compare (1) and (1'):

- (1) I believe John said it.
 (1') I believe John said it.

In each of these sentences, each word has its own particular stress assignment, but one particular word (underlined) has a greater prominence assigned to it than other stressed syllables and the sentence stress has an effect on how the sentence will be interpreted. In some cases, sentence stress may be assigned to syllables which do not receive word stress as in:

- (2) Forty girls and fourteen boys.

Intonation refers to a change in a pitch contour across the duration of a sentence, or other large unit of language. One very obvious use of intonation found in many languages is to use a falling pitch contour for declarative utterances and a rising pitch contour for yes/no questions, as in:

- (3) You know how to get there.
 (3') You know how to get there?

1.3 Phonology: Speech Sounds as a System

No language has all the speech sounds possible in human languages; each language contains a selection of the possible human speech sounds. As such each language has its own pattern of sounds. This study of sound patterns is known as phonology and the speech sounds are known as phonemes. The focus of phonology is to determine the ways in which speech sounds form meaningful systems within languages.

The essential property of phonemes is that they contrast with each other. For example, we can tell that the sounds [f] and [v] represent two phonemes in English because they contrast in words like *fine* and *vine*, which differ only in terms of the voicing of the initial fricative but which have very different meanings. Two words that contrast in meaning and have only one different sound are known as minimal pairs. The following are minimal pairs in English (we transcribe phonemes using slashes / /):

bat – vat	/b/ – /v/
bat – pat	/b/ – /p/
pat – fat	/p/ – /f/
hid – heed	/ɪ/ – /i/
hid – head	/ɪ/ – /ɛ/
head – had	/ɛ/ – /æ/

Where many words contrast by replacing one phoneme we call this a minimal series, as in:

hid – heed – head – had – hard – hod – hoard – hood – who'd
 /ɪ/ – /i/ – /ɛ/ – /æ/ – /a/ – /ɒ/ – /ɔ/ – /u/ – /ʊ/

When we examine the possible minimal pairs and minimal series in a language, we can determine the phonemic inventory in that language: that is the speech sounds which make up the system of that language. The phonemic inventories of languages differ greatly. Some are quite large and others are quite small (see Table 1.4).

If we examine the words of a language closely, we discover that a single phoneme can have a range of different pronunciations. For example, consider

Table 1.4 Phonemic inventories in four languages

	Vowels	Consonants
Hawai'ian	i e a o u	p k ʔ m n ŋ w h l
English (Southern British)	i ɪ ε æ a ə ɜ i ɒ ɔ u ʊ e aɪ eɪ aʊ oɪ ou iə εə ʊə	p b t d g k f v θ ð s z ʃ ʒ h m n ŋ ʃ ʤ w l r j
French	i e a o u y ø ɛ ã õ õ	p b t d k g f v s z ʃ ʒ r m n ɲ w l j
Warlbiri (Australia)	i a u	b ɖ d ɟ g m ɱ n ɲ ŋ ɭ l ɭ ɣ r w y

the following English words (note that [ɭ] and [ɱ] indicate a voiceless [l] and [n]):

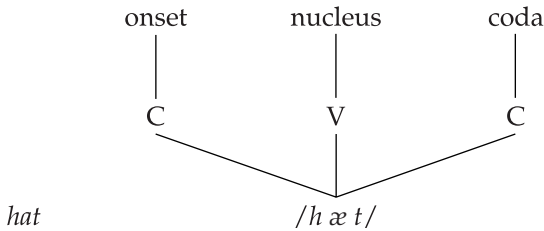
/p/ *pin* [p^hɪn] *spin* [spɪn]
 /l/ *leap* [li:p] *sleep* [sli:p]
 /n/ *knees* [ni:z] *sneeze* [sni:z]
 /h/ *who* [hu:] *huge* [çju:ɖʒ]

In each pair of words, the sound is phonetically different because of the different environment (e.g. /p/ is [p^h] initially but [p] after /s/), but the sounds are still perceived by speakers of English as the same phoneme as there is no meaningful contrast between the sounds, and substituting one for another would not produce a different word, just an unusual pronunciation of the same word. Where two or more sounds represent the same underlying phoneme we call these allophones. It is possible for two languages to have the same sounds but to treat them differently in their phonological system. For example, English and Spanish both have the sounds [d] and [ð], however in English these are two different phonemes (*those* [ðouz] = /ðouz/ and *doze* [douz] = /douz/) while in Spanish they are allophones of the same phoneme: [d] occurs at the beginning of words and after consonants and [ð] occurs between vowels (*Dios* 'God' [diɔs] = /diɔs/ and *adiós* 'good-bye' [aðiɔs] = /aðiɔs/).

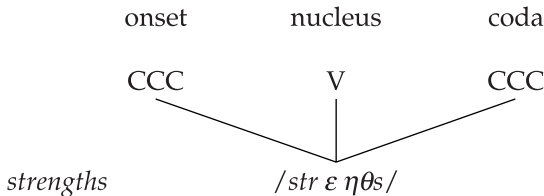
1.3.1 Phonotactics

Just as languages have different phonemic inventories and different allophones, they also have different possibilities for combining sounds into syllables, or different phonotactics. Syllables are phonological units consisting of one or more sounds and are made up of a nucleus (the core of the syllable made up of a highly sonorous segment, usually a vowel), with possibly an onset (a less sonorous segment preceding the nucleus) and/or a coda (a less sonorous segment following the nucleus). The nucleus and coda together are known as the rhyme.

We can see an example of a syllable with all three parts in the English word *hat* which is made up of a single consonant (C) followed by a vowel (V) and then another consonant (C):



All syllables must have a nucleus. Some languages do not allow syllables to have a coda, e.g. Samoan. Other languages allow for more complex syllables with consonant clusters in the onset and possibly in the coda (Blevins, 1995). English allows for quite complex syllables as in:



Languages also have phonotactic constraints on what can occur in a particular position in a syllable. For example, English does allow for CCC onsets, but not any three consonants can occur in this position: */tkf/* would not be possible as the beginning of an English syllable. Different languages have different constraints. Some languages allow for some consonants to be nuclei, e.g. Cantonese *m̂h* */m̂/* 'not', *n̂gh* */ŋ̂/* 'five'. Other languages restrict what can occur in the coda, e.g. Mandarin Chinese allows only */n/* and */ŋ/*. Spanish does not allow */s/* + C clusters in onsets and so words borrowed from English add a vowel to the beginning to change the syllable structure, e.g. *estrés* 'stress'. Some languages allow a much larger range of consonant clusters in onsets, e.g. German *schwach* */ʃvax/* 'weak', *straße* */ʃtra:ɪsə/* 'street', French *pneu* */pnø/* 'tyre'.

1.4 Morphology

Morphology deals with the way in which words are made up of morphemes, the smallest meaningful units of language. If we take a word such as *untied*, it is clear that this word consists of three smaller meaningful pieces, three morphemes: the root *tie*, the prefix *un-* and the suffix *-d*.

Morphemes can be divided up into various crosscutting categories. Morphemes can be lexical like *tie*, with full, complex meanings. Or they can be grammatical morphemes, like *-d*, where a speaker does not really have a choice; the grammar of the language simply requires the morpheme to be present if the action occurred in the past. Morphemes can also be divided into free and bound morphemes. Free morphemes are those which can be used on their own, like *tie*; bound morphemes are those which, like *-d*, have to be attached to another morpheme (symbolized by the hyphen). These two categorizations are independent: we have seen the free lexical morpheme *tie* and the bound grammatical morpheme *-d*, but there are also free grammatical morphemes and bound lexical morphemes. An example of a free grammatical morpheme is the English indefinite article *a*. Bound lexical morphemes are not as common in English as in some other languages; in a language like Spanish, the verb morpheme meaning 'eat' has the form *com-*, but this form never appears without some suffix.

Morphemes can also be talked about in terms of their productivity. Some morphemes are highly productive: the past tense morpheme in English can occur on any verb (although it may have different forms, see below). At the other extreme are completely unproductive morphemes. The most famous is the morpheme *cran-* found in the English word *cranberry*. A cranberry is a type of berry, and we can split the morpheme *berry* off, leaving us with *cran-*, which does not occur anywhere else in English. Other morphemes fall between these extremes of productivity, so that *un-* occurs on some, but not all, verbs (*untie* but **ungo*, where the asterisk indicates an ungrammatical word or sentence); and *-hood* occurs on some, but not all, nouns (*motherhood*, **tablehood*).

A single morpheme may appear with different forms in different words. The words *horses*, *cats*, *dogs*, and *oxen* all have suffixes showing that more than one entity is being talked about, but this plural suffix has different forms, called different allomorphs. Some of these allomorphs are phonologically conditioned, with the form depending on the final phoneme in the root – the form [ɪz] occurs after the sibilant (s-like) sound at the end of *horse*, [z] occurs after the final voiced phoneme at the end of *dog*, and [s] occurs after the voiceless phoneme at the end of *cat*. Sometimes allomorphs are lexically conditioned, the form is exceptional and depends simply on the root – we would expect the plural of *ox* to be *oxes* with [ɪz], but it is not, and speakers simply have to learn this about the word *ox*.

Morphemes can be of different types, as well. So far all the bound grammatical morphemes we have seen have been affixes, where a morpheme is attached in front of a root (a prefix like *un-*) or behind a root (a suffix like *-s*).

There is another rarer type of affix, an infix, where a morpheme is placed within a root. For example, in Chamorro, spoken on the island of Guam, there is a root *chocho* meaning 'eat'. In order to use a verb such as this in a sentence like 'I ate', an infix *-um-* must be placed after the first consonant phoneme, giving *chumocho*. It is not the case here that *ch*, *um* and *ocho* are separate morphemes – by themselves, *ch* and *ocho* do not mean anything. The two morphemes are *chocho* and *-um-*, it is just that *-um-* is placed after the first consonant inside the morpheme with which it combines.

As well as the different types of affixes, a morpheme can be shown by root modification, where the idea of the morpheme is expressed by a change of form in the lexical root. We saw above that there is a plural morpheme in English, usually expressed by a suffix such as *-s*. But the plural of *mouse* is *mice* – plurality is shown by changing the vowel of the root. Sometimes the root is changed completely, a process known as suppletion. The past tense morpheme in English is often expressed with a suffix [t], [d] or [ɪd] (depending on the preceding sound), as in *walk* versus *walked*; it is sometimes expressed through root modification, as in *run* versus *ran*; but in the pair *go* and *went*, the past tense is expressed through suppletion, with a completely different form. Because we tend to think of a morpheme as a thing, it can be hard to think of root modification or suppletion as morphemes, and linguists often talk about affixation and root modification as morphological processes rather than morphemes, but the principle is the same – there are two bits of meaning in *mice*, the bit that shows 'mouseness' and the bit that says there is more than one mouse. A simple morpheme such as a suffix can also be thought of as the morphological process of adding a suffix.

An additional complication arises because sometimes the absence of any material in itself can show a particular idea, and be treated as a morpheme. In English, using the root *book* means we are talking about a particular sort of reading matter. We can use this root with the plural suffix *-s* to indicate that we are talking about more than one of the items. But in a sentence such as *the book is red*, the form *book* does not just indicate the general idea of 'bookness' – the use of the form without the suffix *-s* indicates that we are talking about a single book. That is, the absence of the suffix *-s* indicates an additional concept beyond the general idea of 'book', it shows singular. This use of a contrast between no material and an explicit marker, where either choice shows an additional element of meaning, is sometimes talked about as the presence of a zero morpheme (symbolized with \emptyset). That is, we could say that in *the book is red*, the word *book* actually consists of two morphemes, the lexical root *book* and a singular suffix $-\emptyset$. While 'zero morphemes' are considered inappropriate by many linguists (how do you tell if there's one, two, or sixty-seven zero morphemes in a word?), it is important to realize that the absence of other (explicit) morphemes can be meaningful. Of course, whether a particular absence is meaningful depends on the language. In the Colombian language Awa Pit, like in many languages but unlike in English, the marking of plural is optional. The root *pashpa* means either 'child' or 'children', depending

on context; there is a suffix *-tuzpa* which indicates plurality (*pashpatuzpa* 'children'), but the absence of this suffix does not indicate singular, unlike the absence of the plural suffix in English.

Another morphological process which occurs in some languages is reduplication, which may be full or partial (depending on whether the whole word or only part of the word is reduplicated). For example, *toko* is Indonesian for 'shop', and *toko-toko* means 'shops'. In Ancient Greek, the perfect form of the verb commonly has a partial reduplication of the verb stem, so that the verb root *pau* 'stop' becomes *pepau* (with a repeating of the initial consonant of the root) in a verb form such as *pepau-k-a* 'I have stopped'.

These various morphological processes such as affixation, root modification and reduplication can also be combined in different ways – to form the plural of *child* in English, we add a suffix *-ren* but also change the vowel from the diphthong [aɪ] to [ɪ].

A further morphological process is compounding, where two roots are combined to form a single new word. For example the roots *black* and *bird* can be compounded to form a new word *blackbird* with a different meaning; from *boy* and *friend* we can form *boyfriend*. Some languages have much more productive compounding than English.

Morphological processes are often divided into two types, inflection and derivation, although the distinction is not always clear. Given an English root *consider*, we can make forms like *considers* and *considered*, but also forms like *consideration* and *considerable*. The unsuffixed form and the first two suffixed forms are different forms of the same lexeme – if you want to look *considered* up in a dictionary, you look under *consider*, it's just that if an action happened in the past, the grammar of English forces you to add the inflection *-ed*. On the other hand, *-able* is a derivation, it derives a new lexeme *considerable*, which you would look up by itself in the dictionary. Inflections are highly productive (they apply to all or nearly all roots of a word class), semantically transparent (the meaning of *considered* is 'consider' plus past tense), and do not change word class (*consider* and *considered* are verbs); derivations are not necessarily productive (**goable*), not necessarily semantically transparent (what is the relationship between *consider* and *considerable*?), and may change word class (*considerable* is an adjective).

Languages differ greatly in their use of morphology and the types of morphological processes which they allow. There are two scales that languages are often considered to fall on. One scale is that of isolating, agglutinative, and fusional; the other consists of analytic, synthetic, and polysynthetic. An isolating language is one which does not join morphemes together in one word, agglutination is the process where morphemes join but are easily segmentable (*consider-ed*), and fusion is where morphemes join but are hard to segment (*mice* is 'mouse-plus-plural' but we cannot segment it). An analytic language is one where each word only has one morpheme (and is thus also isolating), a synthetic language has a few morphemes per word, and a polysynthetic language may have many morphemes in a single word. Of course, most

languages have a combination of all of these traits, but these scales are used as an overall heuristic of what is most common in a language.

1.5 Syntax

In English, *the boy sees the girl* means something different from *the girl sees the boy*, and **the the boy girl sees* is not a sentence. Syntax deals with how to put words together to form sentences which mean what we want.

1.5.1 Word classes

The basis of syntax is the fact that the words of a language come in different classes or parts of speech – nouns, verbs, adjectives, prepositions, and so on. Not all languages have the same classes (English has articles like *a* and *the* showing that a noun phrase is indefinite or definite respectively, Japanese does not), and the same basic meaning can be expressed in different classes in different languages (thus the most basic words corresponding to most kinship terms in the Yuma language of California are verbs; to say ‘I am his younger brother’, you say literally something like “he younger-brother-calls me,” where the equivalent of ‘younger-brother-call’ is a single verb morpheme). We establish the word classes and which words are in which class on the basis of the way words behave. For example, in English there is a class of words that take an inflection to show past tense (*walked*, *strolled*, *ran*) and another class which can follow the word *the* at the end of a sentence (*I saw the book/table/boy*).

Having established the word classes for a particular language, we can then label them. There is always a class which contains most of the words referring to concrete objects, and we call that class ‘nouns’. Likewise, there is always a class which contains most of the words referring to actions, and we call that class ‘verbs’. It is important to note that the precise list of words which are in any class may differ from language to language; as we noted above, the word corresponding to ‘brother’ in Yuma is a verb, and while *excitement* is a noun in English, it does not refer to an object. So we cannot say that nouns are words referring to things; rather a noun is any word which is in that class, defined in terms of language-specific behavior, which happens to include most words referring to things (and other words as well).

Many languages also have subclasses within each class. For example, while all verbs in English show marking for tense, they can be distinguished by how many nouns (or arguments) they are associated with. For example, the verb *die* is intransitive, only taking one argument (*Joshua died*, **Joshua died the book*); *kill* is transitive, with two arguments (*Sarah killed Moses*, **Sarah killed*); and *give* is ditransitive, with three arguments (*Ruth gave Abraham the book*).

Nouns and verbs are the only universal word classes (Schachter, 1985). Many languages have a class of adjectives, but in some languages descriptive words have exactly the same behavior as nouns or as verbs, and consequently in

these languages there is no class of adjectives, since there is no special behavior to distinguish them. Different languages have different ways in which their nouns and verbs behave, and so different tests for assigning word class. In English, for example, verbs are marked for tense, but in a language like Indonesian verbs do not inflect for tense, so we cannot use that as a way of establishing the class of verbs in Indonesian (though there are other tests).

It is important to realize that there is no “true” set of features associated with different word classes. As English speakers, we tend to think that the distinction between singular and plural is important, because it shows up in the grammar of English. But there are many languages which do not have this distinction, so when speaking such a language people do not think about it, though they can be more precise if they want to, as English speakers can distinguish *three books* and *four books*, but may just choose to say *books* if the exact number is unimportant. The opposite happens with the English word *we*. When we use this word, we do not indicate if the person we are talking to is part of the group or not, but in many languages there are two separate words translating *we*, depending on whether the addressee is included or not – in Indonesian, for example, *kitu* must be used if the addressee is in the group (inclusive, we-including-you) while *kami* must be used if the addressee is not in the group (exclusive, we-excluding-you). Different languages force their speakers into making different distinctions, with different features being associated with different word classes in different languages.

Despite this, there are some features which are frequently found associated with particular word classes in many different languages. For example, nouns are often marked for number. In English, nouns are either singular or plural; other languages may make more distinctions, so Warlpiri has singular, dual (two) and plural (more than two). And some languages do not mark number at all.

Gender or noun class is another feature commonly associated with nouns. For example, every noun in Spanish is either masculine or feminine, whether human, animate, or inanimate. The gender of a noun affects, for example, the form of the definite article (‘the’) which is used with the noun – *la mujer* ‘the woman’, *el hombre* ‘the man’, *la silla* ‘the chair’, *el libro* ‘the book’. In some languages there are more distinctions than two; Latin has three genders (masculine, feminine, and neuter), while Bantu languages of southern Africa divide their nouns into about ten different ‘genders’ or noun classes.

A further common noun feature is case, where the form of words changes depending on how they are used in a sentence. For example, Latin nouns are marked for case, and thus *puella* and *puellam* both mean ‘girl’. The difference is that the first shows that the word is acting as a subject in the sentence, while the second is acting as an object. This is similar to the distinction between *I* and *me* in English. Some common cases are nominative (primarily used to mark subjects), accusative (objects), dative (recipients), and genitive (possessors). Once again, different languages have different systems of case-marking – English has no cases on nouns, German has four, Latin has six cases, and Finnish has fifteen. Each case may be used for more than one function, so that

in German, for example, the dative is used to show a recipient, but is also used on the noun phrase that follows the preposition *mit* 'with'.

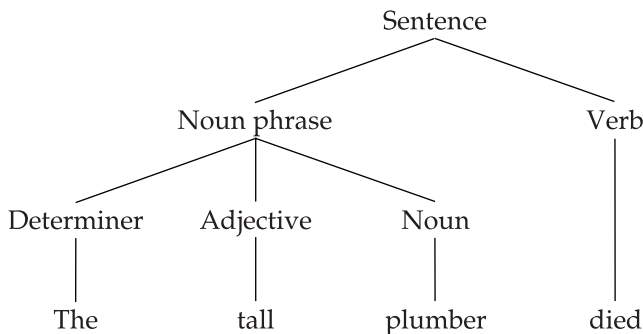
Verbs have a different set of features which are often associated with them. These include tense (the marking of when something happened relative to now), aspect (roughly speaking, whether an event is viewed, for example, as completed or on-going), and modality (expressing something about the reality or otherwise of an event, for example indicative and subjunctive verb forms in languages like French and Spanish). In some languages, verbs agree with their subject or object, a process also known as cross-referencing. For example, in Spanish, the difference between *comí*, *comiste* and *comieron*, all past tense forms of *com-* 'eat', is that the first shows that its subject is first person singular ('I ate'), the second is second person singular ('you (singular) ate'), and the third is third person plural ('they ate').

1.5.2 Constituent structure

In most languages, words are not just strung together in any order. Given the sentence *The tall plumber died*, there is no other way of ordering the words to form an English sentence. Also, at an intuitive level, *the tall plumber* seems to go together as a unit, in a way that *plumber died* does not; then the unit *the tall plumber* goes together with the unit *died* to form the sentence.

There are various ways of showing that *the tall plumber* is a unit, without resorting to intuition. This sequence of words can be substituted by a single word, say *Deborah* or *he*. If the sentence is rearranged in some way, this sequence remains together: *It was the tall plumber who died*. And the sequence of an article or determiner such as *the*, followed by none, one or more adjectives, followed by a noun, turns up again and again in English sentences. Using these sorts of tests, we can show that this sequence forms a constituent. Since the most important word in the constituent is the noun, we call this constituent a noun phrase or NP.

Constituent structure can be represented in different ways. Two common ways are through phrase structure trees and phrase structure rules. Phrase structure trees show the constituent structure of a particular sentence, with all the intermediate constituents.



Phrase structure rules are more general representations of possible sentences. We have seen that a noun phrase can consist of a determiner, one or more adjectives, and a noun, with the determiner and adjectives being optional. We can represent this formally as:

$$\text{NP} \rightarrow (\text{Det}) (\text{Adj})^* \text{N}$$

Here NP is the noun phrase, Det is a determiner, Adj an adjective and N a noun. The parentheses indicate that the element is optional, while the asterisk tells us we can have more than one of this class of word in this position. We can also devise a rule to make our sentence, S, by having

$$\text{S} \rightarrow \text{NP V}$$

where V is a verb. Of course, if we want to include the possibility of an NP after the verb (in a sentence like *The boy saw the girl*), we will have to make the rule more complex:

$$\text{S} \rightarrow \text{NP V (NP)}$$

These rules are clearly not adequate to represent English as a whole, but show the principle of phrase structure rules. Most syntactic theories, such as Government and Binding (Haegeman, 1994), Minimalism (Radford, 1997), Lexical Functional Grammar (Bresnan, 2001), and Role and Reference Grammar (Van Valin & LaPolla, 1997) use some sort of phrase structure rules or trees, although clearly they can be much more complicated than the ones given here.

Different languages have different phrase structure rules (and different trees). For example, in Turkish the verb comes at the end of a transitive sentence, after both NPs, so Turkish would need a phrase structure rule like

$$\text{S} \rightarrow \text{NP (NP) V}$$

In a few languages, these sorts of phrase structure rules do not work very well. In Latin, the words in a sentence can come in almost any order without changing the basic meaning, so phrase structure rules showing where to put each of the words are not much use; but modifications can be made for languages like these.

1.5.3 Semantic roles and grammatical relations

In a sentence like *The farmer is killing the ducklings*, there is a difference in the relationship between the two noun phrases and the verb – we know that the farmer did the killing, and the ducklings ended up dead, and we could talk about them as the ‘killer’ and the ‘thing-killed’. But we know that these are

quite similar semantically to the ‘hitter’ and the ‘thing-hit’ in *The farmer is hitting the ducklings*. For this reason, more general terms are used to express the semantic role (also called the theta role) which a noun phrase plays in a sentence. Different systems of semantic roles are used, but some of the more common terms are agent (the one who performs something, as the farmer above), patient (the one to whom things happen, the ducklings above), experiencer and theme (*I* and *him* respectively in *I saw him*, where I do not really do anything, and nothing actually happens to him), recipient, and source and goal (where something comes from or goes to respectively, as *house* and *shops* in *she left the house for the shops*).

Semantic roles are needed to talk about sentence construction. For example, in English, if a transitive verb has an agent and a patient, the agent comes before the verb and the patient after, which is how we know who does what in *The farmer is killing the ducklings*. If the sentence is made passive (*The ducklings are being killed by the farmer*), then as well as a change in the verb, the patient now comes before the verb, and the agent is either in a prepositional phrase with *by*, or omitted entirely.

On the other hand, we clearly need more than just semantic roles in language descriptions. In the sentences *The farmer is killing the ducklings*, *The ducklings are being killed*, and *I saw him*, there is something in common between the first noun phrase of each sentence, even though they are respectively agent, patient, and experiencer. This noun phrase comes before the verb; if the verb is present tense it controls the form of the verb (e.g., *is* versus *are*); and if the noun phrase consists of a pronoun it has nominative form (*I* rather than *me*). For this reason we need grammatical relations such as subject, object, and indirect object. These grammatical relations are defined in formal terms, so that in English the subject is that argument which comes directly before the verb, has nominative form if it is a pronoun, and controls the verb form. Because grammatical relations are defined formally, different languages may have different sets of grammatical relations. For example, English does not have an indirect object, although some other languages do – in formal terms, *Mary* acts the same way in English in *John kissed Mary* and in *John gave Mary a book*, so it is the same grammatical relation (object) in both sentences; and *Mary* acts the same in *John gave a book to Mary* and *John went with Mary*, so it is the same grammatical relation in both sentences (oblique or object-of-preposition).

There is a relationship between semantic roles and grammatical relations, in that if a transitive verb has an agent and a patient and the verb is not passive, then the agent will be the subject and the patient will be the object; but agent and subject can be distinct (*The ducklings* (subject) *are being killed by the farmer* (agent)), as can patient and object. In some languages grammatical relations may be signaled by constituent order, as in English; in others, constituent order may be free and grammatical relations signaled by case, as in Latin; in others, cross-referencing on the verb may signal the difference. As in English, more than one technique may be used.

Grammatical relations may have more or less importance in the syntax of a language. In particular, in some languages grammatical relations are very important in complex sentences, while in other languages they are not.

1.5.4 Complex sentences

So far all of the sentences considered have consisted of only a single clause. However it is possible to combine more than one clause in a single sentence. The simplest way of doing this is coordination, where two clauses are joined with a word like *and*. Even here there can be important syntactic effects, however. In English, we can say *Rachel saw Judith and left*. The first clause is complete, with a subject (*Rachel*) and an object (*Judith*), but the second clause contains only *left*, which is missing a subject. Clearly, of course, Rachel is the one who left. But we only know this because English has a syntactic rule which says that if two clauses are coordinated, the subject can be left out of the second clause if it is coreferential (refers to the same entity) as the first subject. In other languages, there can be different rules – in a similar sentence in the Australian language Dyirbal, it would be *Judith who left*, as the Dyirbal rule is that a subject can be left out of an intransitive second clause if it is coreferential with the object in the first clause. In other languages, grammatical relations are not important here, and in the equivalent sentence either Rachel or Judith could have left, depending simply on context.

As well as coordination, clauses can also be combined using subordination. This is where one clause (the subordinate clause) is somehow less important than the other (the matrix clause). There are three types of subordination – complementation, relative clauses, and adverbial subordination.

Complement clauses are those clauses which substitute for a noun phrase in a sentence. For example, in English we can say *I saw the boy*, with *the boy* the object of the verb *saw*. But we can also say *I saw (that) the boy left*, *I saw the boy leave* and *I saw the boy leaving*. In each case, where we might expect a noun phrase like *the boy*, we have a whole clause, with at least a subject and a verb. Which type of complement clause we get depends on the verb in the matrix clause, so that with *want* rather than *see*, we can have *I wanted the boy to leave*, but not **I wanted that the boy left* or **I wanted the boy leaving*. With *want* we can also leave the subject of the subordinate clause out if it is coreferential with the matrix clause (*I want to leave*) which we cannot do with *see* (*I saw myself leave* versus **I saw leave*). Different languages have different types of complement clauses, and different rules about which complement clause type goes with which verbs.

Relative clauses add some extra information about a noun phrase in a sentence, and in English often begin with *who*, *which* or *that* – *the man who gave me the book left* contains the relative clause *who gave me the book* (which corresponds to a main clause *the man gave me the book*); this has been added into the sentence *the man left* to specify which man. Different languages differ greatly in how they form their relative clauses. We have seen that one option in English is to

leave the common argument (the noun phrase which occurs in both main clauses, *the man*) out of the relative clause, put *who* in the relative clause, and put the relative clause inside the matrix clause after the common argument. An extremely different process is used in the West African language Bambara:

- (4) tye ye [ne ye so min ye] san
 man PAST I PAST horse which see buy
 'The man bought the horse which I saw'

Here a relative clause based on the sentence *ne ye so ye* 'I saw the horse' has been inserted in the matrix clause *tye ye so san* 'the man bought the horse' in place of *so* 'horse'. The word *min* has been added in the relative clause after the common argument *so* 'horse', which has been left in the relative clause and left out of the matrix clause (the opposite of English).

The third type of subordination, adverbial subordination, covers those subordinate clauses which are similar in use to adverbs – there are a wide variety of possible constructions in languages, corresponding to English clauses such as *because I went, after he came, while working*, and so on.

1.5.5 Sentence types

There are three basic types of sentence: declarative, interrogative, and imperative. For example, in English we have a declarative sentence *He opened the window*, the interrogative *Did he open the window?*, and the imperative *Open the window!* While these sentence types broadly correspond to statements, questions, and commands or suggestions, this correspondence is not complete – for example you could issue a command or suggestion with an interrogative utterance (*Could you open the window?*), or ask a question using declarative word order with questioning intonation (*He opened the window?*). Different languages have different ways of forming these three sentence types, by changes in word order, special verb forms, intonation, or special particles.

1.6 Information Structure

One of the functions of syntax is to structure the ways in which information is presented in sentences and this structure is dependent on the context in which the information is presented. As such, the study of language needs to go beyond the level of isolated sentences and treat sequences of sentences, or texts.

1.6.1 Encoding given and new information

Syntax is often sensitive to whether or not information being conveyed can be expected to be known or not by the addressee (Ward & Birner, 2001). In this

context, we can distinguish between given information – information which the speaker believes is already available to the hearer, or new information – information which the speaker does not expect the hearer to already know. These two types of information are encoded in sentences in different ways. Consider the exchange in (5).

- (5) A: Who took the book?
B: Mary did.

In this example, B's utterance is made up of two pieces of information: 'Mary', which is new information, and 'took the book', which is given information. In this case, 'took the book' is encoded as the pro-verb *did*. Given information is often reduced in such a way. Consider the oddity of (5') as a conversational exchange:

- (5') A: Who took the book?
B: Mary took the book.

Whether information is given or new affects the way in which the information is conventionally introduced into discourse. In English, new information is often introduced in non-subject position, while given information is usually found in subject position. When new information is referred to again in the same discourse, that is when it has become given information, it may be placed in subject position. This can be seen in (6):

- (6) I saw *a really good film* the other day. It was about *a man who thought he was going to be killed by some gangsters*. He went into hiding in the hills, but *they* found him.

In this sentence, there are three NPs which begin as new information, but are later used as given information:

New	Given
<i>a really good film</i> (object)	→ <i>it</i> (subject)
about <i>a man who thought he was going to be killed by some gangsters</i> (object of preposition)	→ <i>he</i> (subject)
by <i>some gangsters</i> (object of preposition)	→ <i>they</i> (subject)

In addition, new information is usually introduced in indefinite NPs (*an X*, *some X*), while subsequent references have definite forms such as definite NPs (*the X*) and pronouns. This can be seen in (6).

Sometimes, information which has not previously been mentioned is introduced in definite NPs, as in (7):

- (7) We went to a restaurant. The waiter was rude but the food was good.

In this case both *waiter* and *food* are mentioned for the first time in the discourse but the use of the definite article (*the waiter, the food*) seems to indicate that they are being treated as given information. Cultural context has a role here – our knowledge of the world tells us that restaurants have waiters and food, so these things are in a sense given information in the light of other knowledge we have from outside the discourse: that is, while the information is new to the discourse it is not new to the hearer (Prince, 1992; Ward & Birner, 2001). Such information can be easily recovered from context and as such speakers can expect addressees to have such information readily available. Therefore it can be treated as given information in such contexts.

In English, the definite and indefinite articles have an important role in the presentation of given and new information, however other syntactic structures are used in other languages. In Russian, for example, word order is related to given and new information (Comrie, 1979). Rather than having SVO word order, Russian usually presents new information late in the sentence, as can be seen in the contrast between (8) and (9).

- (8) *Što koška presleduet?*
 What cat-NOM is chasing
 ‘What is the cat chasing?’

Koška presleduet sobaku.
 cat-NOM is chasing dog-ACC
 ‘The cat is chasing the dog.’

- (9) *Što presleduet sobaku?*
 what is chasing dog-ACC
 ‘What is chasing the dog?’

Sobaku presleduet koška.
 dog-ACC is chasing cat-NOM
 ‘The cat is chasing the dog.’

1.6.2 Topic-comment structure

Another way to view information in utterances is in terms of topic and comment. Topic and comment often overlap with given and new information, however the two sets of terminology involve quite different concepts. The topic of the sentence can be considered the central element in the sentence- the thing the sentence is about – while the comment is what is said about it (Chafe, 1970; Lambrecht, 1994). Consider the exchange in (10):

- (10) A: What did Mary do?
 B: She took the book.

In B, the topic of the sentence is ‘Mary’ (*she*) and the comment, the thing said about Mary, is *took the book*. In this case the topic is given information and the

comment is new information. However sometimes the topic can be new information, as in (11):

- (11) Virginia always eats her vegetables, but her brother only likes ice cream.

In the second part of this sentence, the brother is the topic, but is also new information. By contrast, in (12) the comment is given information.

- (12) Virginia does not like ice cream, but her brother likes it a lot.

In English, the topic is often but not always related to the subject of the sentence (Li & Thompson, 1976; Tomlin, 1983), but there are other structures which can topicalize an NP. Unlike English, some languages use topic as a basic grammatical category. This is the case in Japanese where the postposition *wa* functions as a topic marker, as in (13) and (14), where in each case the topic is a non-subject constituent.

- (13) *Sakana wa tai ga ichiban ii.*
fish TOP bream NOM first good
'Speaking of fish, bream is the best' or 'Bream is the best sort of fish.'

- (14) *Tookyoo kara wa daremo konakatta.*
Tokyo from TOP no-one come-NEG-PAST
'Speaking of coming from Tokyo, no-one did.' or 'No-one came from Tokyo.'

In other languages word order can be used to indicate topics, as in the Chinese sentence in (15) and the French sentence in (16). Here, placing a constituent at the front of a sentence is a way to mark the topic. The French example differs from the Chinese in that the topicalized NP is repeated later in the sentence as a pronoun (*gare* 'station' is feminine, so the pronoun is 'she').

- (15) *Zhè-ge zhǎn lǎn huì wǒ kàn dào hěn duō yóu huà*
this-CLASS exhibition I see very many painting
'At this exhibition, I saw very many paintings.'

- (16) *La gare où est-elle?*
the station where is she
'Where is the station?'

1.7 Semantics

Semantics, that part of linguistic description which deals with meaning, is often divided into lexical semantics, dealing with the meaning of words, and

grammatical semantics, how morpheme meanings are combined by grammar to form the meaning of utterances.

1.7.1 *Lexical semantics*

The form which definitions of words should take is a vexed issue in lexical semantics. Different theories take different positions on what definitions should achieve. Some believe that a definition should be sufficiently precise as to include or exclude any particular case, sometimes with a paraphrase approach based on natural language (e.g., Wierzbicka, 1996) or a specially developed metalanguage (e.g., Jackendoff, 1983). Others believe that the lexicon is not structured in this way, but is rather more often similar to a web of prototypes (e.g., Langacker, 1990) or involving a strong use of metaphor (e.g., Lakoff, 1987).

Theories of meaning also differ in terms of whether or not they distinguish between dictionary knowledge and encyclopedic knowledge (Haiman, 1980; Wierzbicka, 1995). For example, many people in our society know that salt is chemically sodium chloride. The question is whether this is part of the meaning of the word *salt*, to be included in a definition, or simply an additional fact about salt (defined in other ways) which many speakers happen to know.

Another important issue which any general theory of lexical semantics must take into account is that the meanings of a far greater proportion of the lexicon than usually imagined, if not the meanings of all words, are language-specific. While this is obvious for words for cultural artifacts, non-equivalence of word-meanings extends throughout the lexicon. The natural world is not divided up the same between different languages, so that the Japanese word *nezumi* covers a collection of animals which in English would be divided into two types, rats and mice. The human body, a physical universal, is divided up in different ways in different languages: in Spanish, the single word *dedos* is used for both fingers and toes, while Japanese has a single word *ashi* corresponding to English leg and foot. Physical aspects of the world are equally different: English has a color category *blue*, but Russians have two terms covering the same range, *goluboj* (lighter) and *sinij* (darker), and these colors are no more closely related for Russians than *green* and *blue* for speakers of English; speakers of Russian are surprised that English only has one word. Human actions may be more or less differentiated: in English we can *hit* someone, but in many languages different verbs must be used depending on whether the action was hit-with-the-open-hand, hit-with-a-fist, hit-with-a-stick, and so on. All facets of the world and events that take place may be encoded differently – the words of different languages divide the world up differently.

As well as looking at the meanings of words, lexical semantics also examines the meaning relations between words. These meaning relations include concepts such as synonymy (where two words have the same, or at least very similar, meanings, as with *couch* and *sofa*), antonymy (opposite meanings as with *good* and *bad* or *tall* and *short*), hyponymy (the meaning of one is included

in the meaning of another, as with *boy* and *child*), homonymy (two words having the same form but different meanings, as with a *bank* for money and a *bank* of the river), and polysemy (where a word has two or more related but distinguishable meanings, as with a *chip* of wood, a potato *chip*, and a computer *chip*, where all have the idea of a small piece as part of their meaning).

1.7.2 *Grammatical semantics*

Some work in grammatical semantics is interested in the meaning of grammatical morphemes, and how systems of grammatical meaning differ across languages. For example, both English and Spanish show tense using verb suffixes, but English has a single past tense corresponding roughly to two different past tenses in Spanish.

As well as the meaning of individual morphemes (lexical and grammatical), there is also the issue of how these meanings combine to form sentences. Even if we know the meaning of the words *boy*, *girl*, and *kiss*, as well as *the* and *-ed*, there is more to the meaning of the sentence *the boy kissed the girl* than the sum of the meanings of the morphemes, since this sentence means something different from *the girl kissed the boy*, which contains exactly the same morphemes.

One way in which semanticists deal with this issue is through the concept of constructions (Goldberg, 1995). Essentially this approach says that, as speakers of English, we have a schema or template such as Noun Phrase – Verb – Noun Phrase, and we have a meaning assigned to this general schema – say, ‘the first noun phrase has the more active role, the second the more passive role’ – and by combining the meanings of the words with the meaning of the schema, we come up with the meaning of the overall sentence. A different schema would then be used to account for the passive sentence *the girl was kissed by the boy*.

Another approach, Formal Semantics, relies much more on the apparatus of formal logic and grammatical theory. In this approach, the word *kiss* is stored in the lexicon not just with the general meaning of kissing, but with an explicit statement in a formal notation indicating something like ‘this verb’s (underlying) subject is the agent and its (underlying) object is the patient’. The meaning of the sentence is then created by assigning the appropriate semantic role to the appropriate grammatical relation. The meaning of the passive equivalent is created through rules such as ‘make the underlying object into a subject’, ‘make the underlying subject come after the preposition *by*’. Formal Semantics is associated with the idea of truth-conditional or truth-value semantics, which attempts to establish, given a sentence, what conditions have to hold in the real world for the sentence to be true.

1.8 Conclusion

This chapter can only give a brief outline of what is involved in the description of languages and each area we have discussed has a wealth of literature

and a depth of detail which we are unable to address here. However, this brief description should be sufficient to introduce applied linguists to the broad themes and general concepts with which linguists work in developing descriptive accounts of languages.

While language description may not be a core concern for applied linguists, a coherent understanding of the structural features of language is important for applied linguistics research and practice. At all levels of their work, applied linguists must come to grips with language as a system and as such linguistics and language description is basic to applied linguistics work, even if it is not central to the questions which applied linguists pose themselves. We do not claim that linguistic theory is or should be the driving influence in applied linguistics. Rather, we are claiming that a certain level of familiarity with the principles of linguistics provides a framework within which the work of applied linguistics can be carried out in an informed and principled way. The role of linguistics is, therefore, to inform applied linguistics not to determine applied linguistics (cf. Davies, 1999; Widdowson, 2000).

The relationship between language description and applied linguistics is not, however, unidirectional. The insights which applied linguistics gains from confronting real-world language-related problems has great potential to inform the development of linguistic theory and refine our understanding of what needs to be included in language descriptions.

See also 2 LEXICOGRAPHY, 4 LANGUAGE CORPORA, 5 DISCOURSE ANALYSIS, 10 CONVERSATION ANALYSIS.

REFERENCES

- Blevins, J. (1995) The syllable in phonological theory. In J. A. Goldsmith (ed.), *Handbook of phonological theory* (pp. 206–44). Cambridge, MA: Blackwell.
- Bresnan, J. (2001) *Lexical-functional syntax*. Oxford: Blackwell.
- Chafe, W. L. (1970) *Meaning and the structure of language*. Chicago: University of Chicago Press.
- Comrie, B. (1979) Russian. In T. Shopen (ed.), *Languages and their status* (pp. 91–207). Philadelphia: University of Pennsylvania Press.
- Cruttenden, A. (1997) *Intonation*. Cambridge: Cambridge University Press.
- Davies, A. (1999) *An introduction to applied linguistics: from practice to theory*. Edinburgh: The Edinburgh University Press.
- Goldberg, A. E. (1995) *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Haegeman, L. (1994) *Introduction to government and binding theory* (2nd edn.). Oxford: Blackwell.
- Haiman, J. (1980) Dictionaries and encyclopedias. *Lingua*, 50, 329–57.
- Jackendoff, R. (1983) *Semantics and cognition*. Cambridge, MA: MIT Press.
- Kennedy, C. & Miceli, T. (2001) An evaluation of intermediate students'

- approaches to corpus investigation. *Language Learning and Technology*, 5(3), 77–90.
- Lakoff, G. (1987) *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
- Lambrecht, K. (1994) *Information structure and sentence form: topic, focus and mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Langacker, R. W. (1990) *Concept, image, and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.
- Li, C. N. & Thompson, S. (1976) Subject and topic: a new typology of language. In C. N. Li (ed.), *Subject and topic* (pp. 303–34). New York: Academic Press.
- Odlin, T. (1994) Introduction. In T. Odlin (ed.), *Perspectives on pedagogical grammar* (pp. 1–22). Cambridge: Cambridge University Press.
- Prince, E. F. (1992) The ZPG letter: subjects, definiteness and information status. In S. Thompson & W. C. Mann (eds.), *Discourse description: diverse analyses of a fund-raising text* (pp. 295–325). Amsterdam: Benjamins.
- Radford, A. (1997) *Syntax: a minimalist introduction*. Cambridge: Cambridge University Press.
- Schachter, P. (1985) Part-of-speech systems. In T. Shopen (ed.), *Language typology and syntactic description*, vol. 1: *Clause structure* (pp. 3–61). Cambridge: Cambridge University Press.
- Stubbs, M. (1986) *Educational linguistics*. Oxford: Blackwell.
- Tomlin, R. S. (1983) On the interaction of syntactic subject, thematic information and agent in English. *Journal of Pragmatics*, 7, 411–32.
- Tomlin, R. S. (1994) Functional grammars, pedagogical grammars and communicative language teaching. In T. Odlin (ed.), *Perspectives on pedagogical grammar* (pp. 140–78). Cambridge: Cambridge University Press.
- Van Valin, R. D., Jr., & LaPolla, R. J. (1997) *Syntax: structure, meaning and function*. Cambridge: Cambridge University Press.
- Ward, G. & Birner, B. J. (2001) Discourse and information structure. In D. Schiffrin, D. Tannen, & H. H. Hamilton (eds.), *The handbook of discourse analysis* (pp. 119–37). Oxford: Blackwell.
- Widdowson, H. G. (1979) *Explorations in applied linguistics*. Oxford: Oxford University Press.
- Widdowson, H. G. (1980) Applied linguistics: the pursuit of relevance. In R. B. Kaplan (ed.), *On the scope of applied linguistics*. Rowley, MA: Newbury House.
- Widdowson, H. G. (2000) On the limitations of linguistics applied. *Applied Linguistics*, 21, 3–25.
- Wierzbicka, A. (1995) Dictionaries vs encyclopaedias: how to draw the line. In P. W. Davis (ed.), *Alternative linguistics: descriptive and theoretical modes* (pp. 289–315). Amsterdam: John Benjamins.
- Wierzbicka, A. (1996) *Semantics, primes and universals*. Oxford: Oxford University Press.

FURTHER READING

- Akmajian, A., Demers, R. A., Farmer, A. K., & Harnish, R. M. (2001) *Linguistics: an introduction to language and communication* (5th edn.). Cambridge, MA: MIT Press.

- Börjars, K. & Burridge, K. (2001) *Introducing English grammar*. London: Arnold.
- Brinton L. (2000) *The structure of modern English*. Amsterdam: John Benjamins.
- Carnie, A. (2002) *Syntax*. Oxford: Blackwell.
- Carr, P. (1999) *English phonetics and phonology*. Oxford, Blackwell.
- Crowley, T., Lynch, J., Siegel, J., & Piau, J. (1995) *The design of language: an introduction to descriptive linguistics*. Auckland: Longman Paul.
- Fromkin, V. A. et al. (2000) *Linguistics: an introduction to linguistic theory*. Oxford: Blackwell.
- Goddard, C. (1998) *Semantic analysis: a practical introduction*. Oxford: Oxford University Press.
- Gussenhoven, C. & Jacobs, H. (1998) *Understanding phonology*. London: Arnold.
- Hudson, G. (2000) *Essential introductory linguistics*. London: Blackwell.
- Kreidler, C. W. (1998) *Introducing English semantics*. London: Routledge.
- Ladefoged, P. (2000) *A course in phonetics* (4th edn.). Fort Worth: Harcourt College Publishers.
- Ladefoged, P. (2001) *Vowels and consonants: an introduction to the sounds of language*. Oxford: Blackwell.
- Lyons, J. (1995) *Linguistic semantics: an introduction*. Cambridge: Cambridge University Press.
- Matthews, P. H. (1991) *Morphology* (2nd edn.). Cambridge: Cambridge University Press.
- Spencer, A. (1991) *Morphological theory*. Oxford: Blackwell.
- Spencer, A. (1995) *Phonology*. Oxford: Blackwell.
- Thomas, L. (1993) *Beginning syntax*. Oxford: Blackwell.