

the structure of language*

David Crystal

The *theoretical* contribution of linguistics needs discussion, in that unless we can grasp in broad outline a picture of the way in which language is structured, it will be very difficult to find our way about the subject. We need a model of the main branches of the discipline of linguistics as a preliminary to any more detailed study. Figure 1.1 therefore shows one possible model of language structure, which attempts to interrelate the main branches of the discipline.

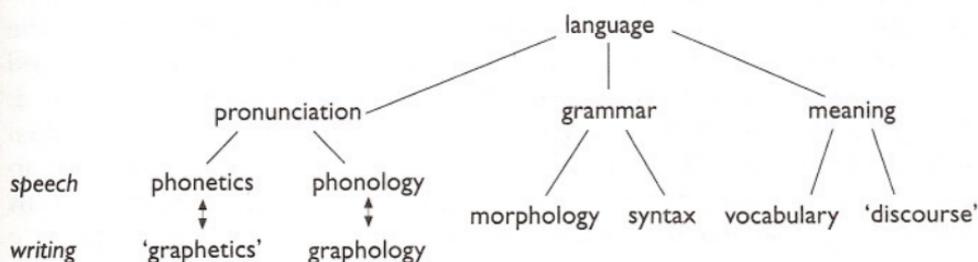


Figure 1.1 Levels of language

There are of course many possible models of the structure of language, and each has its controversial points; but all accounts agree that certain components are essential, and the figure illustrates what these are. For speech, which is in the primary medium of normal human language, three main components, or *levels* of structure are recognised: pronunciation, grammar and meaning. (This is by no means a novel analysis, of course: distinctions of this kind were made by traditional grammarians too.) 'Pronunciation' is, however, too broad a notion to be left as it is. There are two aspects to its study. Firstly, we may study the properties of human soundmaking as such – the way in which we form, transmit and hear sounds. This is the subject of *phonetics*. Apart from certain medical conditions (e.g. cleft palate), all human beings are born with the same vocal apparatus, and in principle can make the same

*This chapter has been taken from the book *Child language, learning and linguistics* (2nd edn) by David Crystal (1987), published by Edward Arnold.

range of sounds. Because of its general applicability, therefore – providing a means of analysing and transcribing the speech of the speakers of any language – the subject is sometimes called ‘general phonetics’. It has to be clearly distinguished from the second term under the heading of pronunciation, *phonology*. Phonology is primarily the study of the sound system of a particular language, such as English or French. Out of the great range of sounds it is possible for each of us to produce, we in fact only use a small set of sounds in our own language – some forty-odd distinctive sound-units, or phonemes, in the case of English, for instance. Whereas phonetics studies pronunciation in general, therefore, phonology studies the pronunciation system of a particular language, aiming ultimately at establishing linguistic principles which will explain the differences and similarities between all such systems.¹

A similar distinction might be made for the written medium, represented further down the diagram. Here we are all familiar with the idea of a language’s spelling and punctuation system. The study of such things, and the analysis of the principles underlying writing systems in general, is equivalent to investigating the phonology of speech, and is sometimes called ‘graphology’ accordingly. Each language has its own graphological system. One might also recognise a subject analogous to phonetics (say, ‘graphetics’) which studied the properties of human mark-making: the range of marks it is possible to make on a range of surfaces using a range of implements, and the way in which these marks are visually perceived. This is hardly a well-defined subject as yet, hence my inverted commas, but it is beginning to be studied: typographers look at some aspects of the problem, as do educational psychologists. From the linguistic point of view, it should be possible to establish a basic alphabet of shapes that could be said to underlie the various alphabets of the world – just as there is a basic international phonetic alphabet of sounds. But this is a field still in its infancy.

On the right of the diagram we see the study of meaning, or ‘semantics’. In a full account, this branch would need many subdivisions, but I will mention only two. The first is the study of the meaning of words, under the heading of ‘vocabulary’, or ‘lexis’. This is the familiar aspect of the study of meaning, as it provides the content of dictionaries. But of course there is far more to meaning than the study of individual words. We may talk about the distribution of meaning in a sentence, a paragraph (topic sentences, for instance), in a chapter, and so on. Such broader aspects of meaning have been little studied in a scientific way, but they need a place in our model of language. I refer to them using the label ‘discourse’ – but as this term is not as universally accepted as the others in my diagram, I have left inverted commas around it.²

Sounds on the left; meanings on the right. ‘Grammar’, in the centre of the model, is appropriately placed, for it has traditionally been viewed as the central, organising principle of language – the way in which sounds and meanings are related. It is often referred to simply as ‘structure’. There are

naturally many conceptions as to how the grammatical basis of a language is best studied; and comparing the various schools of thought (transformational grammar, systemic grammar, and so on) forms much of the content of introductory linguistics courses. But one particularly well-established distinction is that between 'morphology' and 'syntax', and that is presented in the model. Morphology is the study of the structure of words: how they are built up, using roots, prefixes, suffixes, and so on – *nation, national, nationalise* etc., or *walk, walks, walking, walked*. Syntax is the study of the way words work in sequences to form larger linguistic units: phrases, clauses, sentences and beyond. For most linguists, syntax is, in effect, the study of sentence structure; but the syntactic structure of discourse is, also an important topic.³

All schools of thought in linguistics recognise the usefulness of the concepts of pronunciation, grammar and meaning, and the main subdivisions these contain, though they approach their study in different ways. Some insist on the study of meaning before all else, for example; others on the study of grammar first. But the existence of such differences should not blind us to the considerable overlap between them. However, before we can claim that our model is in any sense a complete account of the main branches of language, useful as a perspective for applied language work, we have to insert three further dimensions. These are to take account of the fact of language variation. Any instance of language has a structure represented by the model in Figure 1.1; but over and above this, we have to recognise the existence of different kinds of language being used in different kinds of situation. Basically, there are three types of variation, due to historical, social and psychological factors. These are represented in Figure 1.2. 'Historical linguistics' describes and explains the facts of language change through time, and this provides our model with an extra dimension. But at any point in time, language varies from one social situation to another: there are regional dialects of English, social dialects, and many other styles, as has already been mentioned. 'Sociolinguistics' is the study of the way language varies in relation to social situations, and is becoming an increasingly important part of the subject as a whole. It too requires a separate dimension. And lastly, 'psycholinguistics' is the study of language variation in relation to thinking and to other psychological processes within the individual – in particular, to the way in which

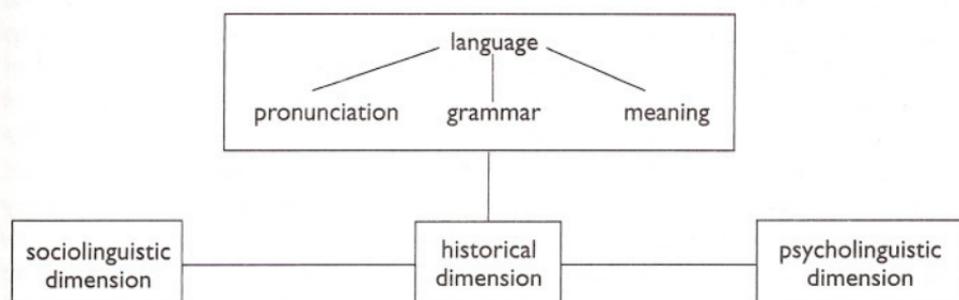


Figure 1.2 Main dimensions of language variation

language development and use is influenced by – or influences – such factors as memory, attention and perception.⁴

At this point any initial perspective has to stop. From now on, we would be involved in a more detailed study of the aims of the various branches outlined, and we would have to investigate further different theoretical conceptions, techniques, terminology and so on. But it should be clear from what has been said so far that in providing a precise and coherent way of identifying and discussing the complex facts of language structure and use, the potential applicability of the subject is very great. What must be remembered in particular is the distinction between (a) the need to get a sense of the subject of language as a whole, and (b) the mastery of a particular model of analysis to aid in a specific analytical or experimental task. The first, crucial step is to develop a linguistic 'state of mind', a way of looking at language that can provide fresh or revealing facts or explanations about the structure and use of language. From here, one proceeds to a more detailed examination of some of the main theoretical principles that underlie any scientific study of language, such as the distinction between historical and non-historical (diachronic v. synchronic) modes of language study, the distinction between language form and language content, and the importance of language variety. In the light of these principles, old problems turn up in a new light, and a certain amount of rethinking about traditional ideas becomes necessary.

Such rethinking can proceed along general or particular lines. The general viewpoint tends to give rise to fierce debate, this is the need to develop greater *tolerance* of language varieties and uses – of other people's accent and dialect, in particular. Can this be done without sacrificing the notion of the 'standard' language, without losing a sense of 'correctness' in language use, and all that many would hold dear? People sometimes accuse linguistics of throwing all standards to the wind – of wanting to say that 'anything goes', that it does not matter how we speak or write, as long as we are intelligible, expressing our ideas, and so on. This is simply not so.

The 'particular' viewpoint can be illustrated here, however, because it shows the kind of detailed thinking that needs to take place in adopting a linguistic way of looking at language. We may take any of the traditional grammatical categories, such as 'number', 'person', 'tense' or 'case' to demonstrate this. Traditionally, it was assumed that there existed a neat one-to-one relationship between the formal category and its meaning, viz. singular = 'one', plural = 'more than one'; 1st person = 'me' or 'us', 2nd person = 'you', 3rd person = 'the other person(s)'; tense = time; genitive case = possession. One of the things that linguistics has tried to do is show how such neat equations do not work. In the person system, for example, we can show this complexity very readily. Taking just one form (the so-called 'first person') we find that the *we* form may refer to the 1st person (as in 'We are going', where it refers to the speaker along with someone else), but it may also be used to refer to the 2nd person (as when a nurse addresses a patient with a 'how are we today?' where

the *we* is equivalent to 'you'), or to the 3rd person (as when one secretary asks another 'how are we today', gesturing at their boss who has just gone into his office).

Another example of unexpected complexity is the tense system. There are many problems in the view that tense expresses time, and that time relationships are expressed by tense forms only. Visualising time as a line,

past time	now	future time
-----------	-----	-------------

it might seem plausible to see the tenses fitting in neatly, as follows:

PAST (<i>I walked/was walking</i>)	PRESENT (<i>I walk/am walking</i>)	FUTURE (<i>I will walk/be walking</i>)
---	---	---

past time	now	future time
-----------	-----	-------------

But there are many examples of usage where this parallelism does not work. For instance, the present tense form *may* refer to present time, as in *I'm leaving*. But it may also help to refer to future time, when used with a future adverbial, as in *I'm going to town tomorrow*; or to past time, when used with a past narrative marker, e.g. *Three weeks ago, I'm walking down this road . . .*; or to habitual action, when used with an adverbial of frequency (*I go to town three times a week*); or to very recent time (as in news headlines, e.g. *Sir X dies*); or to no time at all (the so-called 'timeless' present, as in *Oil floats on water*). Similar examples would arise if we were to consider the other tense forms.

It should be noticed how an examination of just one small area of grammar involves the use of a number of technical terms (even though the meanings of these terms are fairly obvious). It is evident that the need to talk at this level of detail is not far away, as soon as we approach any area of grammar. It is easy to make general, impressionistic remarks about children's usage, for example – about the 'complexity' of their sentences, or about their use of tenses or adjectives or prepositions. But such general remarks need to be carefully watched. And indeed, most of the 'obvious' features of language emerge as hiding considerable complexity, when they are subjected to analysis. Two further examples will illustrate this: the blind use of the idea of 'parts of speech', and an uncritical acceptance of measures of 'length'. The notion of 'parts of speech', for instance, seems easy enough to apply, as long as we are dealing with nouns, verbs, and other central classes. It is less easy to use when classifying such words as *yes, please, sorry, not*, and cases where words have several different uses. And we must remember that definitions of even the central classes can vary greatly from book to book. The notion of 'sentence length' provides another example of hidden complexity. This is a concept

commonly used as a means of plotting language development – sentences get longer as children grow older. But how exactly is length to be measured – in syllables, words, phrases . . . ? If words, then would we count *it's* as one word or two? Would idioms (such as *it's raining cats and dogs*) be counted as having the same number of 'units' as literal sentences (such as *he's keeping cats and dogs*)? And what would we say about two sentences that were equal in length but which differed markedly in complexity (e.g. *The man and the dog and the cat were tired* and *The dog belonging to the man was in the kitchen*)? Such questions immediately arise as soon as we try to work with a simple measure of length.

Last of all – that is, after we have been motivated to accept the *general* aims and tenets of linguistic inquiry – there is the need to find and use a specific *model* of linguistic description, in order to interrelate our various observations about language structure and use. Whatever our aims (whether assessment, screening, remediation, development . . .) the need for a standard descriptive measure is paramount. There is no point in describing child A in terms of one linguistic framework, child B in terms of another, and then hoping to compare the two. (This is similar to – but infinitely more complex than – comparing two objects using two systems of measurement: if one is sixteen centimetres and the other eleven inches, which is the longer?) Likewise, an inventory, or list of 'noticeable' or 'interesting' features in someone's use of language is not an adequate account of it, and may mislead. The dangers of 'selective commentary' are threefold. (a) We tend to notice only what we have been trained to look for, e.g. pronouns, adjectives, tense forms; (b) some of the most important features of language may be omitted because they are not readily noticeable, e.g. variations in word order, elliptical patterns; and (c) an inventory provides no explanation, or sense of underlying pattern – for example, to make a list of features, in which item 13 was the definite article, and item 73 the indefinite article, would obviously be of little value. At some point in any grammar of English, these two items would have to be brought together, because the meaning of the one helps to establish the meaning of the other. And so it is for most areas of grammar. Grammar is not a random collection of features, nor is it learned in this way, and the same applies to other levels of language also. It is the task of the linguistic to define all the variables that make up the language *system*, and say how they relate to each other. Naturally, with such a complex system as language, there is no obvious 'best' way of doing this, and this is why there are so many competing linguistic theories. Each tries to present an explanation of language which models the way in which the system 'works'. Thus one encounters the 'generative' schools of thought associated with Chomsky, those associated with M. A. K. Halliday (first 'scale-and-category' grammar, later 'systemic' grammar), and so on. At some stage students of language have to come to grips with one of them, and learn to use it confidently, in order to provide themselves with a framework in principle consistent and comprehensive for carrying out language tasks. This is no place to argue the merits and demerits

of the different positions: each model has its strengths and insights, each its weaknesses and obscurities. But there is no avoiding this final jump. Without a fund of formal knowledge to back up our general knowledge of linguistic aims and theory, there can be no bridge between beliefs and practice.⁵

-
- 1 See further J. D. O'Connor, *Phonetics* (Penguin 1973).
 - 2 See further F. R. Palmer, *Semantics* (2nd edn, Cambridge University Press 1981).
 - 3 See further F. R. Palmer, *Grammar* (2nd edn, Penguin 1982).
 - 4 Introduction to the more detailed study of these dimensions are P. Trudgill, *Sociolinguistics: an introduction* (2nd edn, Penguin 1983); H. H. Clark and E. V. Clark, *Psychology and language* (Harcourt, Brace, Jovanovich 1977); W. P. Lehmann, *Historical linguistics: an introduction* (Holt, Rinehart and Winston 1962). Specifically on English, see D. Crystal, *The English Language*, Penguin 1988. Language acquisition, within psycholinguistics, is dealt with in D. Crystal, *Listen to your Child* (Penguin, 1986).
 - 5 See further J. Lyons, *Language and linguistics* (Cambridge University Press 1981), D. Crystal, *The Cambridge Encyclopedia of Language* (Cambridge University Press 1987), *Linguistics* (Penguin 2nd edn, 1985).