David Crystal

The Linguistic Identity of English-language Dictionaries of Linguistics

I very clearly recall the moment when I began to appreciate the scale of the task facing any lexicographer of linguistics. During a conference in Toronto in 1969 I had the opportunity of visiting H. A. GLEASON, who showed me the files he had been compiling on linguistic terminology. He was following a standard historical method. There were several boxes of slips, each containing a quotation from a linguistics source, and grouped alphabetically, with terms given a preliminary classification into different senses. Some terms had very few slips; some had large numbers. I asked which were proving to be the most difficult to handle, and GLEASON drew my attention to a file-box containing hundreds of slips, most of which were devoted to entries on the myriad senses of just one term: *form*.

It had not occurred to me before that such an apparently innocuous English word could have generated such polysemy. Three decades on, and I am in no doubt that—at least for English—one of the chief problems facing the linguistics lexicographer (and thus, the student and teacher of the subject) is how to handle the remarkable number of 'everyday' words which have achieved some kind of special status within linguistic science. Of course, all subjects make use of everyday words to some extent—the *charm*, *strangeness*, *up* and *down* of particle physics, for example—but linguistics seems to go in for them in a special way. As an example, Table 1 lists all the items beginning with *F* in A DICTIONARY OF LINGUISTICS AND PHONETICS (CRYSTAL, 1997a), ignoring all combining forms. I have given them a 3-level intuitive grading:

E: a term used in a common, everyday sense (e.g. *face, feed*) or which is very closely related by derivation to such a sense (e.g. *fatherese, fronting*)
A: a term which has a more difficult everyday sense (e.g. *filter, fission*)
T: a term which has only a technical sense (e.g. *factitive, formalism*)

Of the 77 items, only 21 (27 %) are clearly T. By contrast, HOLLIDAY'S DICTIONARY OF PLANT PATHOLOGY (1989), a work of similar length and purpose, has over twice as many—44 (57 %). If this sample is representative, over half the terms in linguistics will turn out to be common everyday items which have been given special senses—and if the A terms are included, this proportion rises to nearly three-quarters. Why should this be?

Linguistics		Plant Pathology		Linguistics		Plant Pathology	
face	E	faba	T	foreign	E	filamentosum	T
factitive	T	fabae	T	forensic	A	filaree	T
factive	T	facultative	T	form	E	filbert	A
factivity	T	fagacearum	T	formal	E	fimbriata	T
fall	E	Fagopyrum	T	formalism	T	fine	E
falling	E	Fagus	T	formalist	T	finger	E
family	E	fairy	E	formalization	T	fir	E
fatherese	E	falcatum	T	formalize	T	fire	E
favourite	E	false	E	formant	T	fireblight	E
feature	E	fan	E	formation	E	fireweed	E
feed	E	farinosa	T	formative	T	Firmicutes	T
feedback	E	farlowii	T	formulaic	A	fisheye	E
feeding	E	fascians	T	fortis	T	fitness	E
feet	E	fasciation	T	fortition	T	five	E
felicity	A	fasting	E	fossilization	A	flaccidum	T
field	E	faullii	T	fossilized	E	flag	E
filled	E	fauna	T	frame	E	flagella	T
filler	E	fawcettii	T	free	E	flavofaciens	T
filter	A	feeding	E	frequency	A	flavus	T
filtered	A	fenaminosulf	T	frequentative	T	flax	E
final	E	fenarimol	T	fricative	T	fleck	E
finite	A	fenuram	T	friction	E	flectens	T
finite-state	T	ferbam	T	frictionless	A	flesh	E
first	E	fern	E	front	E	flies	E
Firthian	T	Ferrisia	T	fronted	E	Florida	E
fis	T	fertilis	T	fronting	E	floury	E
fission	A	fescue	T	frozen	E	flower	E
fixed	E	festuca	T	FSP	T	flowers	E
flap	E	festucae	T	f-structure	T	flowerstand	E
flapped	E	fibre	E	full	E	fluorescens	T
flat	E	fici	T	function	E	fluorescent	A
floating	E	ficus	T	functional	E	fluoride	T
flotation	A	ficuserectae	T	functor	T	fluorotrimazole	T
focal	A	Fieberiella	T	fundamental	A	flutolonil	T
focus	E	field	E	fusion	A	fly	E
folk	E	fig	E	fusional	T	focus	E
foot	E	figwort	T	future	E	foeniculi	T
foreground	E	fijiensis	T	fuzzy	E	foliage	E
foregrounding	T	Fijiviruses	T				

Table 1: Everyday and overtly technical terms compared in two subject dictionaries

The answer must partly be to do with the fact that we encounter words describing the properties and functions of language at a very early age. A surprising number of words to do with vocalization, speech, and language turn up frequently in recordings between caretaker and baby during its first three months of life, such as *cry, talk, say, tell, call, ask, listen, hear, mean, noise, voice*, and *loud*. Drawing the child's attention to the importance of speech is evidently an early parental aim. Not long afterwards children encounter writing-related words such as *read, story, book, turn*, and *page*. And certainly, by the time a child arrives in school the basis of linguistic metalanguage is well-established—so much so that some researchers have used the notion of language awareness as a test for reading readiness : for example, the LARR test (DOWNING, AYERS & SCHAEFER 1983) is one which aims to determine the extent to which young children are familiar with such items as *reading, writing, printing, story, message, address, number, letter, name, top line, bottom line, word*, and *capital letter*. A thesaurus aimed at first-school children (DRYSDALE 1971) contains such items as *language, speech, writing, talk, vowel, consonant, punctuation, dash, comma, syllable, sentence, stress*, and *rhythm*. Book 1 of a recent language and reading programme (PALMER 1994), aimed at children of around age 7–8, has a glossary including such items as *bilingual, caption, contents, emphasis, exclamation, expression, extract, heading, illustration, italic, language, layout, question, sentence, statement, title*, and *underlining*. By Book 4 of this course, aimed at children of 10–11, the glossary contains such items as *accent, agreement, ambiguous, appropriate, conclusion, context, dialect, explicit, formal, informal, introduction, literal, negative, pronounce, rhyme, rhythm, sentence, slang, standard, style, usage*, and *vocabulary*.

It is wisely said that familiarity breeds contempt. And in the present case, a long-established and easy-going familiarity with such items as *exclamation, pronunciation, sentence*, and *word* undoubtedly makes it difficult for anyone to appreciate the need for precision in using these items in a technical context, or to anticipate the scale of the problem—at its worst with such very general notions as *meaning* and *use*. It is not easy to persuade someone who believes that 'everybody knows what a *word* is' of the need for extra terminology (such as *lexeme, lexical item*) to help remove some of its inherent ambiguity.

When words are in unconstrained use in everyday conversation, they develop idiosyncratic senses and connotations, and language-related items are no exception. This is especially noticeable in those academic domains where language has come to play an important role, notably sociology, psychology, philosophy, literary criticism, and (through the area of language pathology) medicine. Terms such as *meaning, style, affective, complexity, code, function, structure*, and *system* are just a few of the many items which have attracted domain-specific applications, and the ramifications of the term *meaning* itself have achieved a classic status within philosophy (OGDEN & RICHARDS 1923). But it is not only the more abstract and abstruse items which attract diverse treatment : reference to the way such terms as *articulation, voice*, and *fluency*

are used in speech pathology, medicine, and psychology also brings to light considerable variation in treatment (CRYSTAL, 1982). And several terms have developed pejorative uses which provide a further barrier between the linguist and the general public, such as *patois, pidgin, argot, jargon, cant,* and *dialect* ('it's just a dialect').

There is, then, an enormous amount of 'lay' terminology in the metalanguage of linguistics. Some dictionaries of the subject have avoided including much of this vocabulary, accordingly, on the grounds that students wishing to get to grips with the terminology of linguistics will not expect to see items there which they would consider to be part of general usage. This is the policy of the first edition of my DICTIONARY OF LINGUISTICS AND PHONETICS, for example :

> one does not need to include such terms as *alphabet*, *abbreviation* and *acronym*, because these are terms whose general sense any good dictionary would handle routinely ; as terms, they owe nothing to the development of ideas in linguistics. Similarly, while such terms as *reinforcement, rhyme-scheme* and *runic* are more obviously technical, their special ranges of application derive from conceptual frameworks other than linguistics.

But of course, what is general today may be specific tomorrow, and by the fourth edition of this dictionary, I was being finding it necessary to include some items (such as *abbreviation*) which had begun to be given a more refined lexical analysis within linguistics. And generally, as a subject develops, one would expect to see a greater technical use made of everyday vocabulary.

When we begin to consider the nature of the more overtly technical metalanguage of linguistics, it quickly becomes apparent that several elements are involved, to a large extent reflecting the ancient history of language study. Terms long associated with traditional Greek and Latin grammar were in due course supplemented by terms from the 19th-century European philological tradition, and these in turn were supplemented (and in many cases supplanted) by terms from 20th-century linguistics. At the same time, linguistic studies developed in different directions within Europe, and also transatlantically. As a result, there is considerable terminological diversity, and this can be seen even if we restrict ourselves to general-purpose terms (as opposed to the terms associated with specific linguistic theories). The terminological variability in the history of linguistic ideas between countries is often underestimated. To take just a few examples of *-ologies* : *speechology*, *phraseology*, *intonology* and *textology* are quite often encountered in the English translations of continental European accounts (e.g. ALEXANDROVA & NAZAROVA, 1988), but are hardly known in Britain. There is considerable scope for comparative lexicological study here.

When we consider traditional terms, particular attention needs to be devoted to their contemporary status. Many would have to be called 'mythical' : they may be listed in some unabridged general dictionaries as 'a technical term for X', but it is doubtful whether they ever had much real life as technical terms, and they are rarely if ever used in modern accounts of language. Several seem to have been invented by individual scholars, who used them in a few publications (thus making it likely that they would be

noticed for inclusion in the OXFORD ENGLISH DICTIONARY, or some other historical work), but whether they ever had any real public currency is unclear. This category includes most of the English translations of classical rhetorical terms, especially those to do with figures of speech, as well as such items as *cacography* (bad handwriting or spelling), *cacology* (unacceptable pronunciation), *heterophemy* (an unintentional error in speech or writing), *heterotopy* (a misplaced sound during speech), *palilology* (word repetition for emphasis), *aptronym* (a name that suits a person's occupation or character), and *paronomasia* (punning). Modern usage of these terms tends to be self-conscious or pedantic. Several of the terms of traditional grammar—especially those with a questionable applicability to uninflected languages (such as the terminology of case)—might be considered to have a similarly uncertain contemporary status.

It is difficult to know how far we should include under the category of mythical the vast amount of nonce-formation which characterizes the field of linguistics (CRYSTAL, in press). To count as a technical term, an item must have achieved some kind of general use, and it is often difficult to know whether this is so. There is a highly creative terminological strand in contemporary phonology, for example, whereby individual scholars give idiosyncratic names to rules identifying particular phonological constraints and conditions, many of which become archaisms almost overnight, as the subject moves on, and other ways are found of handling the phenomena. Much of the early fashionable and self-conscious terminology of generative grammar has since fallen by the wayside in this way, though some of the more imaginative coinages (e.g. ROSS's *squish*) have had a good life. As BOLINGER put it (1975 : 554) :

> Linguistics is an adolescent science that has temporarily outgrown itself. There were no natural checks as there are in physics or aerodynamics, where a mistake may cause a plane to crash or a bridge to collapse. One sign of immaturity is the endless flow of terminology. The critical reader begins to wonder if some strange naming taboo attaches to the terms that a linguist uses, whereby when he dies they must be buried with him. ... With all due allowances for shades of meaning, and granting that any term becomes tinged to a certain extent with the shade of thought with which it is first associated, a tinge is not necessarily a taint, and the production of terms in linguistics is simply ridiculous. For a science about language, linguistics has been peculiarly insensitive to the importance of language in its own development.

As a result, present-day terminology, as seen in a linguistics dictionary, is but the tip of an iceberg. A modern single-volume synchronic work tends to contain about 2,000 entries (e.g. BUSSMANN 1996 ; CRYSTAL 1997a) ; a diachronic work, taking into account the full range of creative linguistic idiosyncrasy, would contain several times that number. The latter, of course, does not exist.

Linguistics is a world of terminological dialects, in which you can identify the theoretical provenance of linguists by the expressions they use—*parole* or *performance, collocation* or *selectional restrictions, variety* or *register*. And not only in a strictly linguistic context : for example, I have often heard HALLIDAY-inspired linguists use *cline* ('a continuum of potentially infinite gradation') in relation to contexts other than language. All subjects have their schools of thought, but the terminological distinctiveness

associated with the proliferation of linguistic theories has often occasioned comment. Linguistics has certainly gone through a phase where it was felt necessary to rename notions in order to avoid the connotations of a previously unpalatable usage. An early illustration is the replacement of *accusative* by *objective* and of *nominative* by *subjective* in talking about pronominal case in English, or the development of *-ing form*, *-ed form*, and the like as an alternative to such Latinate terms as *participle* or *gerund*. Later examples include the different labels given to several parallel notions in the approaches of PIKE, HALLIDAY, and LAMB—such as the names of the various 'levels' of linguistic structure. A specific instance in recent decades is the choice of *nonsegmental* instead of *suprasegmental*, to avoid the implication that this area of phonology is capable of analysis in phonemic terms (CRYSTAL & QUIRK 1964). In such cases, we are dealing with matters of intellectual identity rather than intelligibility.

In some cases, proliferation of terms arises from a group of scholars 'discovering' a linguistic insight, unaware that it has been discovered already. For example, the terminology of nonlinear phonology was already well developed when the writers began to make references to the analogous work of J R FIRTH from some 40 years previously. If that awareness had been available earlier, the contemporary terminological scene in phonology might be very different. Often, the 'rediscovery' relates to a very general notion : a scholar might decide that a dependency relationship between one word class and another should be called *modification*, unaware that precisely the same relationship has previously been called *qualification*, *specification*, or something else. It is a well-recognized characteristic of a developing subject that proponents of a particular theory (whether for reasons of lack of time or inclination) are not well motivated to read outside the boundaries of their own domain. The incestuous growth of generative grammar is a classic case in point : for years, generativists referred only to the work of other generativists, often in unpublished form from within MIT, and it proved difficult for outsiders to obtain an accurate grasp of the newly fashionable terminology. But any personality-motivated school of thought will display this tendency.

Problems always face the historical lexicographer, in such contexts, because proponents of the theory are often inexplicit in their use of terms. When 'everyone knows what they mean' (e.g. at MIT), central theoretical notions are often characterized rather than defined. Even generative grammar, which placed such store on explicitness and formalization, became highly informal when it expounded its approach in ordinary language. Perhaps, when it is known that notions have been given a precise treatment in mathematical or logical terms, there is less motivation to produce definitions which would satisfy the rigorous demands of the lexicographer. Even CHOMSKY himself, ever alert to the dangers of vagueness and obscurity (e.g. Ch. 1, fn. 5 in *Aspects*), and more helpful than most in his provision of definitions, would often phrase his statements with comments that made it impossible to determine exactly what their status was. So often he used rhetoric which threw the weight of interpretation onto the context, but where it proved difficult for the linguistic lexicographer to find an explicit

statement. Here are a couple of examples taken from the opening chapter of *Aspects* (CHOMSKY, 1965):

> For the purposes of this discussion, let us use the term 'acceptable' to refer to utterances that are perfectly natural and immediately comprehensible without paper-and-pencil analysis, and in no way bizarre or outlandish. ... For present purposes, it is unnecessary to delimit it more carefully. (p. 10)
>
> I shall use the following terminology, with occasional revisions as the discussion proceeds (p. 17)

This did no harm to the reader grappling with *Aspects* for the first time: such pragmatic 'half-truths' can actually be pedagogically helpful, and all educationists use this kind of style in expounding their subject. But as source material for historical lexicography they present problems, and insofar as their context-bound status permitted (indeed, encouraged) interpretive flexibility, the rhetoric undoubtedly led to a proliferation of slightly differing senses for terms, in which an orthodox 'core' of information was accompanied by layers of discrepant material. Such fuzzy definitions are a major characteristic of linguistic metalanguage—an irony, for a subject that claims to be proud of its scientific standing.

The lack of an agreed terminology for even the central areas of linguistic enquiry is of course a major difficulty facing the applied linguist, attempting to facilitate progress in such domains as clinical or educational linguistics. Indeed, it is one of the main reasons for the growth in linguistics dictionaries as a genre, in recent years. When a teacher has to choose between *grammar* (= phonology + syntax + semantics) vs *grammar* (= syntax + morphology), or *syntax* (= sentence structure + word structure) vs *syntax* (= sentence structure, not word structure), there is plainly a major issue here for those involved in language education, and lexicographical accounts are invaluable in helping people get to grips with the subject. But the problem is more deep-rooted than one of choice between competing theoretical perspectives. There seems to be a profound suspicion of linguistic terminology, amounting at times to a real fear of the terms (as opposed to the concepts), which other academic domains do not seem to share. A recent event is illustrative.

In 1995 I was asked to write an introduction to English grammar suitable for the middle years of British secondary schools, relating to the demands of Key Stages 3 and 4 of the National Curriculum in English (CRYSTAL 1996). I had given the chapters titles which reflected the way the terms were being used within the body of the text (see column 1 of Table 2). The first version was submitted in due course, and circulated by the publisher to various teachers and curriculum advisers. After this consultation was over, the publisher suggested (I quote from her letter) that I 'look again at the headings so that any "difficult" terminology can be made more recognisable for the average teacher. Technical terms can be sub-headings'. It transpired that the problem was not the use of the terms as such, but the prominence being given to the terms in the chapter headings. It was not being suggested that the teachers were incapable of coping with the concepts involved, or even that the concepts were too difficult for this level of

child: it was simply that, if I persisted in using the 'difficult' terms as chapter headings, it was felt they would put teachers off when they glanced at the table of contents. They would think the book was more difficult than it was, and would keep away from it. Column 2 of Table 2 lists the alternative headings I devised, all of which proved to be acceptable. The irony is that no-one suggested that I needed to remove or replace any of the 'difficult' terms from the body of the text. So, for example, the chapter originally headed *Auxiliary verbs* was replaced by the heading *Helping verbs*, but the term *auxiliary* is introduced in the first line of the first paragraph, and turns up in every sub-heading on the double-page spread. But I had to keep *auxiliary* out of the chapter heading at all costs! There is no better example, I believe, of the irrational fear of linguistics terminology.

Original chapter headings	**Revised chapter headings**
Complements	Completing a clause
Adverbials	Saying when, where and how
Gender	People, animals, and things
Genitive	Apostrophe *s*
Premodification	Adding words before a noun
Postmodification	Adding words after a noun
Auxiliary verbs	Helping verbs
Tense	The time of an action
Aspect	Completing an action

Table 2: Original and revised chapter headings in CRYSTAL (1996)

What is so special about linguistic terms that prompts people to react in this way? It cannot be just a matter of 'long words', as teachers have to deal with *chromatography* and *compressibility* in Key Stage 3 science, and *geomorphology* and *sustainability* in Key Stage 3 geography. The reaction has two explanations—one diachronic, one synchronic. The diachronic explanation must be the uncertainty arising out of the arbitrariness with which terminology was taught in school, in the period from the mid-18th to the mid-20th centuries. The mechanical way in which pupils were taught to identify parts of speech, metrical patterns, figures of speech, and other features of language, using a descriptive system originally devised for a quite different type of language (Latin or Greek), must have led to considerable conceptual confusion, and a readiness to accept what one was told without question, for fear of punishment. This was compounded by the sense of inferiority conveyed by the prescriptive approach to language, where everyday usage was penalized and a fear of 'being incorrect' established which would never thereafter be lost. As an English teacher, now nearing retirement, once said to me, 'I know that there are nine kinds of adverb clause [she then listed a set of names], but I don't know why. Nor do I think I would recognize one in real life, and I certainly couldn't teach children about them'. Moreover, this kind of fear, or uncer-

tainty—an inevitable consequence of an original lack of understanding—is contagious, even between generations. Younger teachers, who were never taught grammar systematically, are often reluctant to question the use of terminology, whether informed or not, by their senior colleagues. And it only takes senior business managers to criticise a subordinate for failing to conform to a prescriptive usage (such as the 'ban' on split infinitives) to ignite the same feelings of confusion, inferiority and fear as they themselves encountered the generation before.

But there is a second, synchronic explanation for the fear of linguistics terminology, which is less to do with attitudes to language and more with the nature of the object language itself. The fact of the matter is that language is complex, for several well-recognized reasons :

- Spoken language is difficult to access. Speech is the 'indispensable foundation' of language study (cf. HENDERSON 1971), but it needs a trained ear and adequate transcriptional skills to be able to describe it well—and certain aspects (notably, the prosodic and paralinguistic dimension) are difficult areas even for specialists.
- There is a great deal of indeterminacy in a language, especially in speech. Notions such as sentence are not always easy to identify ; word classes do not have discrete boundaries ; there may be a mismatch between formal and functional criteria. How much and what kind of language does a given term apply to ? The question is at the heart of contemporary discussion of the use of terms as diverse as *adverb*, *paralanguage*, and *discourse*. Nor does it follow (as we have already seen in the case of Latin and English) that a term found adequate for the description of one language will be automatically applicable to the description of another.
- The analysis of language requires the recognition of several levels of abstraction—at least the essential dualisms introduced by SAUSSURE, often multiple levels of underlying representation or dimensions of grammatical structure, all of which may need to be kept terminologically distinct. The kind of thinking which allows one to say that a sentence may be 'simple' in its syntactic structure at a higher level while being 'complex' at a lower level (as in the notion of a relative clause) is undoubtedly something which neophytes to the subject find difficult.
- The academic perspectives vary greatly, so that a notion viewed from within, say, a psychologist's perspective may differ greatly from the same notion viewed from a linguist's. *Psycholinguistics* is not the same as *psychology of language*, not is *sociolinguistics* identical to the *sociology of language*. A term like *pragmatics* may have significantly different interpretations between academic domains.

For these reasons, it is inevitable that many linguistic terms will present special difficulties to the student, unlike those found in other academic domains, and it is perhaps the unique complexity of linguistic terminology which has motivated such a diverse range of publications within the field of linguistic lexicography.

Several genres exist, depending on the extent to which the publications attempt to come to terms with the range of difficulties identified above. In relation to coverage, works range from the all-inclusive (e.g. CRYSTAL 1997a) to the relatively specialized (e.g. TRASK 1993), and from the pure to the applied (e.g. RICHARDS, PLATT & WEBER 1985). Some restrict themselves to abstract terms within linguistics, allowing in only those proper names which have developed common meanings, such as CHOMSKYan and FIRTHian (CRYSTAL 1997a) ; others routinely include the names of

languages (e.g. BUSSMANN, 1996) or of prominent linguists (e.g. CRYSTAL 1993). Most contain some information about abbreviations and special symbols. Terminological lists can also be found within some larger works, such as the glossary at the end of BRIGHT (1992), ASHER, et al (1993), or CRYSTAL (1997b). However, the provision of a glossary is regrettably not a standard practice in English linguistics writing.

In relation to treatment, works vary in the extent to which they incorporate elements of traditional lexicography. Some routinely give pronunciation and word class for all words, including the 'everyday' ones (e.g. RICHARDS, PLATT & WEBER 1985) ; some give a pronunciation only for those words which might pose problems (e.g. TRASK 1993 ; CRYSTAL 1993). Some give etymologies (e.g. BUSSMANN 1996). All devise their own systems of cross-reference to indicate which words in an entry are given lexicographic treatment elsewhere in the work, and to identify related entries. Works vary in terms of the extent to which they treat related terms within a single entry : most of the entries in HARTMANN & STORK (1972) and TRASK (1996), for example, are single-topic entries, and are thus relatively short. CRYSTAL, by contrast, has several lengthy entries in which sets of related terms are brought together—over a dozen in the case of *semantics*, for example. None routinely give corpus citations, though very occasionally there will be a quotation.

There are also great variations within the textual structure of an entry, though all make use of some combination of the following seven elements :

1 an abstract definition
2 an illustration of its use from the reader's language
3 an illustration of its use from other languages
4 an amplification of the definition or some of its terms
5 an account of its historical provenance or current theoretical status
6 an evaluation of its significance
7 a list of historical sources or corpus citations

Using this framework, it proves possible to begin making some systematic comparisons between different works, and also to begin distinguishing lexicographic genres. (In the following quotations, the typographic distinctions indicating cross-references have been ignored.)

The glossary genre makes minimal use of these criteria, typically restricting entries to criterion 1, and using a single-sentence style, with the occasional addition of a language illustration (criteria 2/3). This style is especially popular within books aimed at introductory students, and is widespread in texts used in primary or secondary schools. Examples include :

> **adjective** An **adjective** is a word which goes with a noun and tells us something else about it. *Example : Poppy was a **hairy scary** troll.* (PALMER 1994 : 60)
> **expression Expression** is the way we show what we feel by using our face or voice. If we put some expression into our voice, we say things with more feeling. (ibid. : 61)
> **colloquial language** This is informal language of the sort we use in conversation rather than in formal writing. (FULLER, JOYNER & MEADEN 1990 : 128)

metalanguage This word describes the various terms that you need in order to be able to discuss language itself. (ibid.: 129)
euphemism This is a word or expression that you substitute for one that people might find offensive or upsetting, e.g. *departed* for *dead*. (ibid.: 128)
affricate (phonet) Said of a consonant in which a complete closure in the vocal tract is gradually released ([pf] Ger. *pfennig*). (CRYSTAL 1997b: 420)
contraction The phonological reduction of a linguistic form so that it comes to be attached to an adjacent form (e.g. English *I'm*), or the fusion of a sequence of forms so that they appear as a single form (French *du*, from *de le*). (BRIGHT 1992: 288)

Books which call themselves dictionaries (of languages or linguistics) are more diverse. They usually have to be considered encyclopedic dictionaries, in the sense that either (a) they add extensively to criteria 2/3 or (b) they do not restrict themselves to criteria 1–3. TRASK (1993) provides an example (details of pronunciation and word class are omitted). Many of his entries show extensive use of criteria 2/3:

discourse item Any lexical item or grammatical form which typically serves to relate one utterance to another in a discourse, or to relate the utterance in a particular way to the discourse as a whole. English has a number of such items, including *well, yes, surely, however, on the contrary, so* and *nevertheless*. Mandarin Chinese is reported as using a grammaticalized set of sentence-final particles for discourse purposes, including *le* 'currently relevant state' *ne* 'response to question', *ba* 'soliciting agreement', *ou* 'friendly warning', *a* or *ya* 'reduced forcefulness' and *ma* 'question'.

Many other entries add information in relation to some of criteria 4–7. The entry on *squish*, for instance, illustrates criteria 1 (the opening sentence), 2 (the second sentence), 4 (the third sentence), and 7 (the source).

squish A continuum of (especially lexical) category membership, by which membership of a category is regarded as a matter of degree, rather than as an either/or proposition. For example, the word *newspaper* in *This newspaper headline* might be regarded as lying somewhere on a noun-adjective continuum. Squishes are fundamental in **fuzzy grammar**; they represent much the same notion as **clines**. (ROSS 1972b).

However, TRASK is a mixed-format work, in that it contains several glossary-type entries alongside these fuller treatments, and in this respect it more resembles the style of a conventional dictionary:

distribution The full range of environments in which a lexical or grammatical form can occur.
ditransitive verb A verb which subcategorizes for two objects, such as *give*: *She gave me a kiss.*

One way of comparing linguistic dictionaries is in terms of the extent to which they are predominantly *definitional* (criterion 1 only), *illustrative* (criteria 1–3 only) or *discursive* (implementing criteria 4–7). Table 3 compares the way four dictionaries treat entries under letter E (ignoring abbreviations, cross-references, and names of languages).

	HARTMANN & STORK (1992)	TRASK (1993)	CRYSTAL (1997a)	BUSSMANN (1996)
Definition only (criterion 1)	24	13	4	15
Definition + illustration (criteria 1–3)	26	25	8	11
Use of amplification (criteria 4–7)	20	34	70	65
Total entries (letter E)	**70**	**72**	**82**	**91**

Table 3 : Types of entry compared in four linguistics dictionaries

HARTMANN & STORK and TRASK make most use of entries involving criteria 1–3 (50 % and 52 % respectively) ; CRYSTAL and BUSSMANN make least use of such entries (15 % and 29 % respectively). The latter would therefore be more accurately described as 'encyclopedic dictionaries', where the term *encyclopedic* refers to treatment, not coverage, and CRYSTAL (1993) in fact makes this explicit in his title : AN ENCYCLOPEDIC DICTIONARY OF LANGUAGE AND LANGUAGES.

Length of treatment in relation to criteria 4–7 is probably the only real difference between encyclopedic dictionaries and (alphabetically organized) encyclopedias. For example, voice quality is given about 50 words in HARTMANN & STORK (1972), about 70 words in TRASK (1996), and about 140 words in CRYSTAL (1997a), but about 800 words in the main article in BRIGHT (1992). By comparison with definitional and illustrative entries, most of the length of the entries in an encyclopedic dictionary entry relates to criterion 4. The author explains a notion further by clarifying the terms in the definition (usually by instantiating them or by relating them to other terms in the dictionary), or by providing methodological information about the way the notion would be applied in linguistic analysis, as the following two examples show. The first is from BUSSMANN (1996) :

> **face-to-face interaction** Communicative behaviour in speech situations where the speaker and listener make immediate contact. Research into face-to-face interaction considers linguistic features, but is primarily concerned with non-linguistic features like facial expression, eye contact, gestures, posture as well as paralinguistic features like manner of articulation (whispering, shouting).

In the second, from CRYSTAL (1997a), everything bar the first sentence is criterion 4, and it is easy to see that this kind of exposition could continue indefinitely—as it would, for example, if the term were being introduced in a journal article..

> **unified features** A term used to characterize models of non-linear phonology which integrate consonantal and vowel place features in a single framework. In this approach, for example, labial and coronal articulations are brought together into a single coronal tier. However, consonants and vowels retain their identity, in that place features of consonants are immediately dominated by the consonantal place node, and vowels by the vocalic place node. Thus the notion of [labial] in a consonant to C-place defines a different plane from that of [labial] in a vowel to V-place.

Criterion 5 adds information about the linguistic context in which a term is used. This can be a simple 'in phonetics' or 'in semantics', or something much more specific, such as 'This analysis is accepted in certain derivational theories of grammar, notably RG and many versions of GB' (TRASK 1996, re the *unaccusative hypothesis*). Authors vary in relation to where they place this context: CRYSTAL invariably makes it his opening sentence, often adding amplification later in the entry; TRASK usually specifies any theoretical context *after* the opening definition. Some entries are largely taken up with criterion 5, as in this entry from BUSSMANN (1996):

deep structure A term from transformational grammar, developed by N. Chomsky, to describe the underlying structure of a linguistic utterance. Deep structure specifies the grammatical relations and functions of the syntactic elements, as well as the linguistic meaning of the elements of a sentence which contain the lexemes, the information important for the execution of transformations. The idea of a difference between two levels of structure in language (deep structure vs surface structure) has a long and complex history and can be found in the writings of the Indian grammarian Panini (fourth century BC), in the seventeenth-century grammar of Port Royal, and in the writings of Humboldt, Wittgenstein, and Hockett. In transformational grammar both structural levels can be represented by tree diagrams. In Chomsky's (1965) aspects model, meaning-neutral transformations mediate between the basic tree structure of the deep structure and the derived tree structure of the surface structure, so that the syntactic structure can be interpreted phonetically. This syntactically motivated concept began a great debate between the supporters of CHOMSKY and the advocates of generative semantics, who regarded the basic structure as semantic. In the various revisions of the standard theory, the level relevant for semantic interpretation was also changed, the structural information of the deep structure being encoded into the surface structure (now S-structure). In this way, the semantic information remains at S-structure, which has been the input for the semantic interpretation since the Revised Extended Standard Theory.

Criterion 6 adds an evaluative or critical dimension to an entry, provided by the author, as can be seen from the final sentence in each of the following illustrations from TRASK (1993):

donkey sentences A sentence containing an anaphor whose antecedent is an indefinite NP whose reference is constrained by a quantifier elsewhere in the sentence; the classic example is *Any man who owns a donkey$_i$ beats it$_i$*. Such sentences pose unusually severe problems for most accounts of anaphora.

discourse A connected series of utterances by one or more speakers. Certain types of grammatical, lexical and phonological elements can be identified which typically serve to relate one utterance to another in some fashion, but, on the whole, attempts at extending the methods of grammatical analysis to the study of discourse, in the hope of constructing 'discourse grammars', have not been outstandingly successful.

It is unusual to find lexicographers being antipathetic towards a term (or a use of a term), given the descriptive frame of reference in which they operate, but judicious criticism will often be encountered in relation to terms or meanings which fall outside the domain of contemporary linguistics, as these examples from CRYSTAL (1997a) show:

syntax ..... The adjective form of 'syntax' in modern linguistics is 'syntactic', as in the above examples; 'syntactical' these days sounds quaint.

continuous A term used in grammatical description of verb forms, referring to a contrast of a temporal or a durative kind, and thus handled sometimes under the heading of tense and sometimes under aspect. The usual contrast recognized is between 'continuous' or progressive (e.g. *I am going*) and **non-continuous**, simple, or non-progressive (e.g. *I go*). Linguists prefer an aspectual analysis here, because of the complex interaction of durational, completive and temporal features of meaning involved; traditional grammars, however, merely refer to 'continuous tense', etc., and thus imply a meaning which is to some degree an oversimplification.

TRASK (e.g. 1996) goes further in at times adding a critical comment about a term introduced by a linguist. This kind of comment is increasingly found in more specialized and advanced dictionaries, where a greater degree of critical awareness might be presupposed:

vocoid A synonym for vowel in the phonetic sense of that term (sense 1), introduced in an effort to remove the ambiguity between the phonetic and phonological senses of 'vowel'. While possibly useful, the term has never become established. (PIKE 1943).

With reference to criterion (7), none of the available dictionaries give corpus citations, as we have already noted, and they differ greatly in respect of the kind and number of bibliographic references. TRASK gives some 400 references, chiefly primary sources; HARTMANN & STORK use a similar number, chiefly secondary sources. BUSSMANN includes bibliographies of varying lengths, a mixture of primary and secondary sources—nearly 200 items referenced in relation to the entries in letter E alone, for example, along with many cross-references to items in other entries. CRYSTAL (1997a) gives very few primary sources for the use of a term, and although he gave a secondary source from a small selection of textbooks for every entry in the first three editions of his dictionary, this feature was dropped from the fourth edition.

Table 4 gives some examples of entries analysed in terms of these seven criteria.

	Criterion
A A single-element entry (cf. glossary, above)	
cohortative (BUSSMANN, 1996)	
Mood of admonition, encouragement, or recommendation, which can be part of either verbal mood or sentential mood.	1
B Two-element entries	
homophone (HARTMANN & STORK, 1972)	
One of two or more words which are identical in sound but different in meaning and/or spelling	1
e.g. English *heir* and *air*.	2
egressive (RICHARDS, PLATT & WEBER, 1985)	
of speech sounds which are produced with air from the lungs moving out through the mouth and/or nose.	1
Most speech sounds in most languages are egressive.	4

	Criterion
C Three-element entries	
impersonal verb (HARTMANN & STORK, 1972)	
A verb used only in the third person singular with no reference to a particular subject	1
e.g. *It is raining*,	2
or French *Il faut* … 'it is necessary'.	3
kinetic (CRYSTAL, 1997a)	
A term sometimes used in phonology	5
applied to tones which vary in pitch range ;	1
also called 'dynamic' or 'contour' tones, and contrasted with 'static' or 'level' tones.	4
D Four-element entries	
case (TRASK, 1993)	
A distinctive, overtly marked form which can be assumed by an NP to indicate that that NP bears some identifiable grammatical or semantic relation to the rest of the sentence.	1
In English, overt case marking is confined to a few pronouns (*I/me ; they/them*),	2
but some other languages, such as German, Russian, Latin, Basque and Finnish, exhibit elaborate case systems typically involving about three to six distinct forms, but sometimes a dozen or more.	3
Among the most frequently distinguished cases are the nominative, accusative, absolutive, ergative, dative, genitive, instrumental, comitative, locative, allative and ablative, but many others exist.	4
E Five-element entries	
null element (TRASK, 1993)	
An element which, in some particular description, is posited as existing at a certain point in a structure even though there is no overt phonetic material present to represent it.	1
For example, many linguists would argue that the plural form *sheep* consists of two mophemes, the noun stem *sheep* and a null plural suffix Ø.	2
Most approaches to grammatical description recognize the existence of at least some null elements, and the problem of how to limit their use in a principled way has long been a bone of contention : both structuralist (sense 2) and generative approaches have sometimes indulged in the proliferation of zeros, even to the extent of positing zeros which contrast with other zeros ;	5
see, for example, Chomsy and Lasnik (1977).	7
As an example of the possible utility of null elements, consider … [130 words]	4
The extensive use of null elements is particularly characteristic of GB ; see empty category The Indian grammarian Panini was the first to use null elements, but their modern use derives primarily from Bloomfield.	5

Table 4 : Levels of complexity in English linguistics dictionaries

I have not yet encountered an entry containing six or seven elements, although it is perfectly possible to conceive of one, using the following schematic :

This is what the term means.
It can be illustrated from these English examples.

It can also be illustrated from these other languages.
It raises these issues.
It is particularly used in this theory.
It has this value to the subject.
These authors have used it or discussed it.

Most entries, of course, do not need all seven : the descriptive terms of grammar and phonetics, for example, can usually be handled with just two or three criteria : 1, 2/3, and 4. Theoretical notions need more weight, but it is interesting to note that, within the dictionaries cited in this paper, entries do not go beyond five. I can understand the absence of seven-element entries : this might simply be because the authors have chosen to illustrate a term only from their own language, in the first instance, and use other languages only when their language does not make use of the notion (e.g. non-English case forms, such as *translative* or *abessive*). Criteria 2 and 3, in effect, thus come to be in complementary distribution. If this is so, it is a limitation which should be avoided in future linguistics lexicography, given the subject's interest in all languages. More surprisingly, I have not yet found an example of the use of six criteria, in these works, though I have only looked at a small sample of entries. The norm seems to be four or five. In many cases, it would be easy to reach six by adding source references or authorial comment. The former stratagem, however, requires a great deal of historical lexicographic research, which is currently lacking. The latter requires considerable authorial experience and confidence, and a readiness to be fair-minded about theoretical positions other than one's own.

I have restricted this paper to those dictionaries which have attempted to deal with the core terminology of linguistics (including phonetics) in English. It would be possible to carry out a similar analysis of the textual structure of other genres of English linguistic lexicography, such as the biographical dictionary ('portraits of linguists') or the dictionary of languages, where fresh issues would undoubtedly emerge. In a languages dictionary, for example, several obligatory elements would have to be recognized, such as family, where spoken, number of speakers, writing system, and chief dialects. Dictionaries of applied linguistics are also likely to raise fresh considerations, because the intended users come from a wide range of professional backgrounds.

I conclude with an anecdote which illustrates that nothing is sacrosanct, when it comes to the compilation of lexicographic works. In a biographical exercise of some years ago, it was decided to restrict the work to those linguists who had passed away, on the grounds that it would be difficult to constrain the size of the book if living linguists were allowed in, there being so many. But, it was pointed out, this would mean that CHOMSKY would not be included, and would a book on the world's linguists be credible if he were absent ? An editorial decision was made. For the purposes of the volume, CHOMSKY would be considered to be 'honorarily dead'.

Literature

ALEXANDROVA, O. & NAZAROVA, T. 1988. *Methods in linguistic research*. Moscow : Moscow University.

ASHER, R.E., *et al*. 1993. THE ENCYCLOPEDIA OF LANGUAGE AND LINGUISTICS. Oxford : Pergamon.

BOLINGER, D. 1977 *Aspects of language*. 2nd edition. New York : Harcourt Brace Jovanovich.

BRIGHT, W. (ed.) 1992. INTERNATIONAL ENCYCLOPEDIA OF LINGUISTICS. New York & Oxford : Oxford University Press.

BUSSMANN, H. 1996. ROUTLEDGE DICTIONARY OF LANGUAGE AND LINGUISTICS. (Translation by G. P. TRAUTH & K. KAZZAZI of LEXICON DER SPRACHWISSENSCHAFT. Stuttgart : Kröner Verlag, 1990.) London : Routledge.

CHOMSKY, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA : M.I.T. Press.

CRYSTAL, D. 1982. Terms, time, and teeth. *British Journal of Disorders of Communication* 17 (1), 3–19.

— 1993. AN ENCYCLOPEDIC DICTIONARY OF LANGUAGE AND LANGUAGES. Oxford : Blackwell/ Harmondsworth : Penguin.

— 1996. *Discover grammar*. London : Longman.

— 1997a. A DICTIONARY OF LINGUISTICS AND PHONETICS. 4th edition. Oxford : Blackwell.

— 1997b. THE CAMBRIDGE ENCYCLOPEDIA OF LANGUAGE. 2nd edition. Cambridge : Cambridge University Press.

— in press. Investigating nonceness : lexical innovation and lexicographic coverage. In: *Story, words and wisdom*. Rutgers University Press.

CRYSTAL, D. & QUIRK, R. 1964. *Systems of prosodic and paralinguistic features in English*. The Hague : Mouton.

DOWNING, J., AYERS, D. & SCHAEFER, B. 1983. *Linguistic awareness in reading readiness (LARR) test*. Windsor : NFER-Nelson.

DRYSDALE, P. 1971. *Words to use*. Toronto : Gage Educational.

FULLER, S., JOYNER, J. & MEADEN, D. 1990. *Language file*. London : Longman/BBC Enterprises.

HARTMANN, R. R. K. & STORK, F. C. 1972. DICTIONARY OF LANGUAGE AND LINGUISTICS. New York : Wiley.

HENDERSON, E. 1971. *The indispensable foundation : a selection from the writings of Henry Sweet*. London : Oxford University Press.

HOLLIDAY, P. 1989. A DICTIONARY OF PLANT PATHOLOGY. Cambridge : Cambridge University Press.

OGDEN, C. K. & RICHARDS, I. 1923. *The meaning of meaning*. London : Routledge and Kegan Paul.

PALMER, S. 1994. *Language Book 1*. (The Longman Book Project.) London : Longman.

RICHARDS, J., PLATT, J. & WEBER, H. 1985. LONGMAN DICTIONARY OF APPLIED LINGUISTICS. 2nd edition. (2nd edition, 1992.) London : Longman.

TRASK, R. L. 1993. A DICTIONARY OF GRAMMATICAL TERMS IN LINGUISTICS. London : Routledge.

— 1996. A DICTIONARY OF PHONETICS AND PHONOLOGY. London : Routledge.