Fred W. Riggs

LEXICOGRAPHICAL TERMINOLOGY: SOME OBSERVATIONS

Introduction

Are lexicographers satisfied with the vocabulary available to
them for use when they are discussing the technical problems
that come up in the preparation of a dictionary? If they are,
should they be? Can they communicate readily enough to meet their
needs - between countries, publishing houses, academic circles?
If the answer to any of these questions is No, then perhaps it
would be interesting and relevant to elicit the help of terminolo-
gists to take a look at the special problems of 'lexicographic
terminology'. Can some useful contributions be made by Terminology
to help solve some of the problems of Lexicography?

As a starting point in our quest for an answer to these
questions, consider the extent (or non-extent?) of overlap between
the vocabulary given in the index to Hartmann (1983) and in
Robinson (forthcoming). It appears that 36 entry terms are found
in both of these works - but they constitute less than one-fifth
of the 199 entries in Hartmann's index, and less than one-fourth
of the 153 entries in Robinson's glossary. Gold (1981) lists
94 technical terms found in a recent volume of essays on lexi-
cographic problems and suggests that it would be useful to supply
definitions for them. We will find, I believe, such definitions
for only ten of these terms in either Hartmann or Robinson.

Terminologists habitually employ a perspective that reverses
the one normally used by lexicographers: instead of proceeding
from words to investigate their meanings (a semasiological method)
they start with the definitions of concepts and then ask what
terms can most conveniently designate them (an onomasiological
viewpoint). We may well start by offering a few illustrations
to suggest how this orientation can be put to work to help lexi-
cographers solve some of their own terminological problems. Our
examples can be given under three headings:

(1)   problems of ambiguity;

(2)   problems of terminological overabundance;

(3)   the problem of new concepts.

The problem of ambiguity

In general, a term-form is likely to prove ambiguous when
it has more than one technical sense within the same subject
field. So long as the various senses of a polyseme occur in differ-
ent contexts that can readily be distinguished from each other,
the problem of ambiguity need not arise. Lexicographers necessarily
confront this phenomenon every time they prepare a dictionary
entry in which more than one sense of an entry word is identified
and defined. Let us say that a term-form is 'equivocal' to the

degree that it is used for more than one technical concept within the same subject field, e.g. a field like 'lexicography'. Although equivocal terms can be used unambiguously, provided their contexts of use clearly indicate which of their possible meanings is intended, would it not be helpful if, for each of the senses of an equivocal term one could also use, synonymously, an unequivocal technical term?

To illustrate this problem concretely, let us consider the several meanings of word as this term-form is used by lexicographers. Obviously many of the non-lexicographic senses of word as reported in WEBSTER'S THIRD INTERNATIONAL DICTIONARY (W3) are irrelevant to the present discussion and we may, therefore, ignore them. To simplify this presentation further, I shall restrict myself to the use of data taken from W3, and from Robinson's draft glossary of lexicographic terms. Consider, then, the following four entries quoted from Robinson:

E_1: pronunciation key a table of symbols used to represent speech sounds and corresponding common words containing those speech sounds

E_2: canonical form the form of a word that is widely used and grammatically simple

E_3: multiword lexical unit a lexical unit consisting of two or more words which function syntactically as one unit

E_4: guide word ... a word printed at the top of a page in a dictionary to indicate the entries included on that page

It may well be said that, because word is so familiar, lexicographers need not define it: at least no entry for word is contained in Robinson's glossary. Alternatively, it may be argued that the analysis of the meanings of word belongs to linguistics rather than to lexicography, and so we can safely disregard it.

However, it seems clear that in each of the lexicographic definitions provided in E_1 to E_4 word is used in a very specific technical sense. The point becomes clear enough when we replace this equivocal term-form with an unequivocal term for the concept required to make sense of the definition. Let us take up each usage in turn.

Word means Word-form. The nearest definition of the sense of word found in E_1 that I can find in W3 may be the following:

E_5: word ... (2a-1) a linguistic form that is a minimum free form

In this sense, sing, sang, sung, and singing are different 'words'. To the degree that linguists have technical terms for such concepts as this, lexicographers are surely entitled to use them. In this case, I find that Lyons (1977:19) uses word-form as an unequivocal term for this concept. Would not the sense of E_1 become more evident if the definition were to be rewritten to read:

"... corresponding common word-forms containing those speech sounds"? For example, the vowel sounds in sing and sang are obviously different.

Word means Word-set. The sense of word as it is used in E_2 appears to coincide with the following definition found in W3:

> E_6: word ... (2a-2) the entire set of linguistic forms produced by combining a single base with various inflectional elements...

In this sense, sing, sang, sung, and singing are all different forms of just one word. May we, again, find linguistic terms that can unequivocally designate this concept? Turning to Lyons (1977:19) I find that three terms have been used, paradigm, lexeme and vocabulary word. Perhaps lexicographers will be content to appropriate one of them, but the discussion in Lyons suggests that a number of subtle distinctions are involved, so much so that each of these term-forms could easily be used ambiguously. However, this is not the place to discuss the other meanings that can be attributed to these three term-forms.

Let us suppose, just for illustrative purposes, that lexicographers were to find, after careful study, that no one of these three forms could be used without ambiguity for the sense of word needed in E_2. They might then decide that it would be convenient to coin a new term for the concept. I do not want to propose such a term here, but to illustrate the procedure let us consider a form like *word-set*. (Double asterisks are used to show that this is not a recognized or established term-form.) In support of this proposal, one might argue, however, that *word-set* parallels 'word-form' and is convenient and easily remembered. The definition in E_2 might, accordingly, be re-written to say: "the form of a word-set that is widely used..."

Word as Graphic Word. The third sense of word, as it is used in E_3, is no doubt the most common. Its dictionary definition is given in W3 as follows:

> E_7: word ... (2b) any segment of written or printed discourse ordinarily appearing between spaces...

Interestingly there appears to be no term in linguistics that unequivocally designates this idea. The expression 'phonological word' appears to be used by some linguists, but it is clearly a misnomer since it is not the isolation of a phonic unit, but the separation of a written form that is in question. R.R.K. Hartmann, in a personal communication, has called my attention to the following terms that have been used by some lexicologists for this concept: 'grammatical word', 'structure word', 'lexical word', 'orthographic word' and 'graphic word'. Terminologists tend to use 'orthographic word'. However, a more convenient form might be 'graphic word'. Since the definition of this concept rests on spelling conventions, it seems advantageous to call attention to this feature rather than to incidental properties - grammatical, structural, lexical - that are scarcely decisive.

If one were to use 'graphic word', the definition given in E3 could be re-written in quite unambiguous form: "...two or more graphic words which function syntactically as one unit."

Word as Lexical Unit. The fourth sense of word found in Robinson's glossary, as illustrated at E_4, does not appear in the entry for word given in W3. To determine more precisely what lexicographers understand by a word in this sense we may take a look at the guide words presented at the top of the pages of W3. Here we discover such forms as these: close and cloud, close shot and cluster clover, etio- and -etic, clinico- and -dendrons.

While the concept in question is not identified by W3 as one of the meanings of word when this word is used as a free form, it appears as a bound form in such words as 'headword' 'catchword', and, indeed, 'entry word' and 'guide word'. Moreover, the relevant sense is given in W3 as the meaning of lexeme.

E_8: lexeme ... a meaningful speech form that is an item in the vocabulary of a language

Here, one might suppose, is an unambiguous term-form that can be used to designate the concept of 'word' entailed in E_4. Unfortunately for lexicographers, it is not easy to use this word unambiguously because linguists have appropriated it for concepts not covered by the definition quoted in E_8. For example, as noted above at E_6, Lyons uses lexeme in a sense that is similar, if not identical, to the concept of a 'word-set' as given in E_2 and defined in E_6. To overcome the resulting ambiguity, a new term has been introduced by lexicographers, namely: lexical unit. This usage is reflected in Robinson's glossary where we find the following two entries:

E_9: lexeme ... LEXICAL UNIT

E_{10}: lexical unit ... a word or phrase regarded as a single item in the vocabulary of a language. Also called LEXICAL ITEM, LEXEME

If one were to re-write E_4 with these definitions in mind, one could say "a lexical unit (lexeme) printed at the top of a page in a dictionary...". Interestingly, it is not only in their professional writings that lexicographers use word in the sense of a lexical unit (lexeme), but also even the editors of W3 sometimes – though rarely – use word for this meaning. Consider, for example:

E_{11}: loanword ... a word [i.e., lexeme, lexical unit, lexical item] taken from another language and at least partly naturalized

Since the concept of a 'lexical unit' is, surely, the most basic general concept in use by lexicographers, one might suppose that it would be extremely useful if a very convenient short form – such, for example, as word – could be used unequivocally in this sense. If we could control the practices of lexicographers, we might want to say that in technical writing, word would always

be used only for the sense of a 'lexeme' (lexical unit) as defined in E_8. Recognizing that this would not be acceptable, one might wish that lexeme itself be used, as a convenient graphic word in preference to the more cumbersome open compound, lexical unit. However, if it appeared that the new linguistic senses of lexeme prevented the unambiguous use of this form to mean E_8, lexicographers might contemplate the adoption of a completely new coinage - for example *lexism*. The question, I suppose, is whether lexicographers would rather designate the most basic generic concept in their field by terms that are either equivocal or cumbersome or, by contrast, they could accept a neologism that is both unequivocal and convenient but, of course, so far unheard of.

The problem of terminological overabundance

A good example of a concept for which we have a superabundance of terms is identified in E_3, where it is called a multiword lexical unit. At least 18 terms for this concept have been found in Zgusta's Manual (1971) and other sources, including several personal communications from Edward Gates, David Gold, and others. All such terms can be listed in a single terminological record, as follows:

E_{12}: 'multiword lexical unit' a lexical unit (lexeme) consisting of two or more graphic words which function syntactically as one unit. Also called BOUND SYNTAGMA, COMPLEX LEXEME, COMPOUND LEXICAL UNIT, COMPOUND LEXICAL ITEM, CONJUNCT, FIXED PHRASE, LEXEME CLUSTER, LEXICALIZED PHRASE, MULTIWORD LEXEME, MULTIWORD LEXICAL UNIT, MLU, OPEN COMPOUND, POLYLEXEME, SET COLLOCATION, SET PHRASE, SYNTHEME; 'IDIOM', 'LEXEME', 'LOCUTION'.

The three final term-forms in quotation marks are obviously equivocal - they can well signify other concepts used in lexicography. Even if we eliminate them, we still have some 15 unequivocal terms that can be used for this concept. The problem of superabundance is evidently not unique in English. According to R. Kocourek (1979:123), at least 25 different terms can be found in the French literature to signify the same concept. On the basis of an impressionistic analysis of several pages in W3, it appears that at least a fifth of the entry words in this dictionary consist of 'multiword lexical units' (bound syntagmas, fixed phrases, open compounds). To discuss the criteria that qualify particular phrases to be handled as syntactic units, we may assume that the editors of W3 must have made frequent use of this concept. Nevertheless, so far as I can discover, not one of these technical synonyms for an important concept in lexicography has been quoted as an entry word in W3.

Let us consider next another example, taken directly from Robinson's glossary:

E_{13}: entry word ... a lexical unit ... which heads an entry and is explained in that entry. Also called ENTRY, ENTRY-HEAD, GUIDE WORD, HEADWORD, KEY WORD, LEMMA, MAIN ENTRY, WORD-ENTRY

entry ...(1) ENTRY WORD (2) a paragraph...

entry-head ... ENTRY WORD

guide word ...(1) ENTRY WORD(2) a word printed at the top of a page ...

headword ... ENTRY WORD

key word, keyword ... ENTRY WORD

lemma ... ENTRY WORD

main entry ...(1) ENTRY WORD (2) an entry containing full explanation and...

word-entry ... ENTRY

Four of these terms are also found as entry words in W3, but the treatment there is a little different, as the following entries indicate:

E_{14}:headword ... (1) a word or term often in distinctive type placed at the beginning of ... entry (as in a dictionary...)

E_{15}:entry word ... HEADWORD

E_{16}: entry ... (5b-3) HEADWORD

E_{17}: lemma ... (3) a word or phrase glossed in a glossary

If an important criterion used in the selection of terms for a concept is that of equivocation, then we can see right away that some of the terms used by lexicographers for this basic concept could be dropped because they are equivocal: notably entry, guide word, and main entry, each of which, according to Robinson's findings, also designates other concepts used in lexicography. Even with these deletions, we still have five technical and unequivocal synonyms for this concept. However, further analysis may suggest some more deletions.

For example, as lemma is defined in W3, it is a polyseme, and the most relevant sense differs somewhat from what we mean by an 'entry word'. Could we say that a lemma, as defined in W3, is one type of entry word, but not all entry words are lemmas: for example would not word-entry be an entry word, but not a lemma? Such doubts might suggest that lemma is rather shadowy when used as a technical synonym for entry word.

How about headword? Interestingly W3 refers its users from entry word to this term-form, even though it is a polyseme, and entry word is not. But could it be argued that the other sense of headword is not relevant to lexicography? Let us consider its definition:

E_{18}: headword ... (2) a [graphic] word qualified by a modifier

Most multiword lexical units probably take the form of a

modified 'headword' — in this second sense of the lexeme. Thus word in entry word is a headword — but in the other sense of 'headword', entry word as a whole may be used as a headword. The selection of inverted forms of a bound syntagma involves alphabetizing their headwords (2), in preference to using their uninverted forms as headwords (1). Because of the possible confusions involved in this usage, it is significant that Robinson chose to use entry word rather than headword as her lemma — i.e., as the technical synonym for this concept to be glossed in her glossary.

May we assume that the remaining synonyms — entry-head, key word, and word-entry — are all fully acceptable as unequivocal forms? I would like to express some reservations about two of them. Key word, for example, is widely used by information scientists to designate an index term that is part of the natural language of an author, by contrast with descriptors that have been authorized for use in a given index language. Although we may assume that the vocabulary of indexers and the vocabulary of lexicographers can be kept separate, there is probably enough overlap between these fields to cause confusion. Such overlaps generate 'shadows', even when a usage is not clearly equivocal.

As for word-entry, the form may be used to make a distinction with term-entries. In ordinary dictionaries, as all lexicographers know, a single entry word is followed by as many definitions as may be useful to identify the various senses of the headword. By contrast, in a terminologically oriented glossary, each sense of a given term-form may be quoted in a separate entry article. In the ASTM Compilation (p.298) for example, we find seven 'term-entries' all headed by the single word-form, grade, but each describing a different sense of this word as it is used in a distinct sub-field of 'testing and materials'. When we recognize that many 'terms' — as defined in W3 to identify a lexeme that has only one meaning within a given context of use — can take a single form, then we will understand why the authors of the Compilation chose to quote each of the senses of grade in a separate term-entry rather than condense them in a single word-entry. So as to permit reference to this lexicographically useful distinction, I suggest that the term-form word-entry may be shadowy and hence possibly ambiguous in lexicographic usage.

Finally, we may consider entry-head to be unequivocal and useful, especially if we want to use entry mainly in the sense of an 'entry article'. Then the head [word] of such an article is, clearly, the entry-head. Since I see no real advantage in the effort to select just one of the technical synonyms for a concept as its 'preferred term', I would not want to offer a preference as between entry word and entry-head. Both can unequivocally designate the concept defined in E_{13}. No doubt the other terms listed there can also be used unambiguously as technical synonyms for the same concept, provided that on each occasion of use the context clearly shows which of their possible meanings is intended.

The problem of new concepts

It sometimes happens that a new concept appears to be useful

in a given field of knowledge. Such concepts can, of course, be well identified by their definitions - or, to put it differently, the writing of a definition generates a new concept. If the concept identified by a definition is really new, there cannot yet be a term for it since, assuredly, if a term for the concept already exists, it cannot truly be a new concept. Of course once we have a new concept in mind, provided we want to use it more or less frequently, we may decide to coin a term for it. A procedure for suggesting such new terms and marking them with double asterisks was offered above, e.g. in our discussion of *word-set*. However, in that example we only wanted to find unequivocal terms for established concepts. Now, by contrast, we want to find out how to name new concepts.

For a simple illustration of the problem, consider the following concept:

E_{19}: <u>???</u> a set of term-forms that may be used as defining equivalents of each other.

In such a terminological record the symbol ??? is used to indicate that no term for this concept is available - something that would be true if the concept is indeed new. Before thinking about what to call this concept we might ask whether or not it would have any value for lexicographers. To answer this question, think about the following sentence:

The forms <u>entry</u>, <u>entry word</u>, <u>headword</u>, <u>entry-head</u>, <u>guide word</u>, <u>headword</u>, <u>key word</u>, <u>lemma</u>, <u>main entry</u> and <u>word-entry</u> are all members of a single E_{19}.

No doubt the sentence can be completed by repeating the full defining text found in E_{19}. However, if we wanted to use this concept often, would it not be convenient to have a simple term-form to designate it? Such a term-form might, for example, be *term-set*. It would be easy enough to remember this form by pairing it with *word-set*: different members of a word-set make up the vocabulary entries in a single dictionary entry; the different members of a *term-set* are quoted as entry words in two or more entries that share a single sense, as illustrated in E_{12}.

If we might, provisionally, accept *term-set* for E_{19}, we could use it to define two more concepts:

E_{20}: <u>???</u> a member of a term-set that is defined by an equivalent

E_{21}: <u>???</u> an equivalent used to define other members of a term-set

Entry word, as quoted in E_{13} is a good example of E_{21}. All of the 'also called' term-forms listed in E_{13} are defined at their alphabetical entries by means of the equivalent term <u>entry word</u>. All of them, therefore, have the defining properties identified in E_{20}. If this sounds confusing, it is! However, if we had convenient terms for concepts E_{20} and E_{21}, we could say the same thing much more easily.

We could, perhaps, find a more appropriate term if we were willing to coin one. Consider, to begin with, the sense of 'canonical' as it is used in this entry taken from Robinson's glossary (see also E_2 above):

E_{22}: <u>canonical form</u> ... the form of a word-set that <u>is widely used</u> and usually grammatically simple...

Could we think of a member of a term-set E_{19} that is widely used and (preferably) unequivocal as being 'canonical'? Ideally such a term would head the entry that contains its definition. Moreover, the same term would also be the most appropriate one to use as an equivalent in the entries for all other members of the same term-set, in their alphabetically arranged entries. An example is <u>entry word</u> as quoted in E_{13}.

If we accepted this notion, would it not suggest a good term for E_{20}, i.e. *canonical term* - and for its antonym, i.e. *non-canonical term*? These new terms permit us to re-write the sentence that immediately follows E_{21} as follows: <u>Entry word</u>, as quoted in E_{13}, is a good example of a *canonical term*. It is also used as an equivalent to define all the *non-canonical* members of this term-set at their respective entries.

Based on these relationships, we see the need for another closely linked concept:

E_{23}: <u>???</u> entry headed by a canonical term

Is there any term now in use by lexicographers that unequivocally designates this concept? If so, I have not found it. Let us, again, presume to suggest a possible term-form that could be used, namely *canonical entry*. This form would permit us to say that the *canonical term* for a given concept heads a *canonical entry*, just as the canonical form of a word-set serves as the entry word for its entry, i.e. its base word.

Of course we should not accept the use of 'canonical term' in this sense without first considering possible alternatives. One might be based on the following term already used by lexicographers, according to Robinson's glossary:

E_{24}: <u>index entry</u> ... an entry consisting of a variant entry <u>word</u> and a cross-reference to the entry containing the full description of the canonical form.

An obvious antonym to <u>index entry</u> would be *indexed entry* and Robinson's definition clearly indicates that it would contain a 'canonical form'. However, such a <u>form</u> is not necessarily a canonical <u>term</u> E_{21}. For example, when one has a set of orthographically or morphologically variant forms, e.g. <u>rime</u>: <u>rhyme</u>, or <u>sang</u>: <u>sing</u>. It is necessary to create more than one entry for the same word-set. An 'index entry' refers users from a non-canonical orthographic form to a canonical form found in the same word-set.

But may one also say that an 'index entry' refers users from each non-canonical term belonging to a term-set to the 'indexed

entry' for its canonical term? If so, then the definition given in E_{24} needs to be broadened so as to include sets of synonymous terms (a semantic relation) in addition to sets of orthographic forms (a structural relation). However, if lexicographers agree that the sense of 'index entry' ought not to be broadened from the specifications given in E_{24}, then we could not use it for the following, parallel, concept:

E_{25}: ??? entry headed by a non-canonical term

Would it be acceptable to use the term-form *non-canonical entry* for $E25$? I believe this would be unacceptable because the same form could also be used for an entry composed in an incorrect or unauthorized format. Accepting this objection, could we think of some other term that might be more appropriate? One possibility arises from the fact that every E_{25} is defined by an equivalent. Could we, then, use *equivalent entry* for the concept defined in E_{25}? If so, we could then easily write such rules as the following:

When preparing entries for members of a *term-set*:

(1) the *canonical term* should head and also be defined in the *canonical entry*;

(2) the *non-canonical terms* should be defined by an equivalent, i.e. the *canonical term*, thereby constituting *equivalent entries*; and

(3) all of the *non-canonical terms* should be cross-referenced in the *canonical entry*.

No doubt, at first, such a sentence, full of new technical terms, will seem to be difficult to understand - but as soon as these terms are learned, one will find, I believe, that the sentences are quite clear and they prescribe useful guidelines for lexicographers. Of course, the same rules can be written in non-technical language, but they will be much longer - as one can easily discover by trying to re-write the rules precisely, replacing all the starred terms with familiar vocabulary. Robinson seems to use 'main entry' (which also means 'entry word') to designate an entry with a full explanation and cross-references. *Canonical entries*, of course, take this form. However, a 'main entry' is usually contrasted with the 'sub-entries' that immediately follow it to identify run-in and run-on entries. In other words, the main/sub entry distinction relates to basic forms and their derivatives, compounds and idioms, rather than to the different members of a *term-set*; hence it identifies the separate paragraphs of a single entry article rather than a set of cross-referenced entries for the same concept.

Conclusion

Of course, it would be presumptuous for a terminologist who is not a lexicographer by profession to recommend any such neologisms as those starred above. Only a group of practising lexicographers has the authority to make and, of course, to adopt, such innovations. It is, however, appropriate for a terminologist to offer such examples as these in order to demonstrate the

feasibility and probable utility of using terminological (onomasio-
logical) methods to develop technical vocabulary available to
lexicographers.

Having made this point, I must now immediately point out
that it is not enough to coin new terms, no matter how useful
they may be in theory. Actually it is not that difficult to think
of new term-forms for useful concepts. What is far more difficult
involves the promulgation of a strategy that will assure adequate
evaluation and adoption of suggested innovations.

An appropriate strategy to develop the special language (tech-
nolect) of lexicography needs to be multi-dimensional, utilizing
a 'non-standardizing' approach based on voluntary cooperation
and appealing to the natural interest of lexicographers in the
professionalization of their work. At least four basic levels
or dimensions of such a programme can be identified:

(1) active promotion of scholarly papers and publications
on terminological issues;

(2) use of a 'terminology column' in a respected newsletter
for the profession;

(3) publication, in successive editions, of a dictionary
of lexicographic terms;

(4) launching, in confidential draft form only, a conceptual
glossary for lexicographers.

To implement a strategy such as the one outlined above it
is necessary to have an organized group established under the
aegis of a professional society. The Dictionary Society of North
America has recently launched a Commission on Lexicographic
Terminology and the new European Association for Lexicography
has expressed a willingnes to cooperate. Perhaps, therefore,
the time has now come to start a serious effort to analyze and
promote the technical vocabulary (special language) available
to lexicographers.

References

Gold, D.L. (1980-81) "Review of Papers on Lexicography in Honor of
Warren N. Cordell" Dictionaries 2-3: 175-179
Hartmann, R.R.K. (1983) "Terminological index" in Lexicography:
Principles and Practice ed. by R.R.K. Hartmann. London-New York:
Academic Press
Kocourek, R. (1979) "Lexical phrases in terminology" Travaux de
terminologie 1: 123-153
Lyons, J. (1977) Semantics. Cambridge: U.P.
Riggs, F.W. (1982) "COCTA-glossaries: the *ana-semantic* perspect-
ive" in Proceedings of the Conference on Conceptual and Termin-
ological Analysis in the Social Sciences ed. by F.W. Riggs.
Frankfurt: Indeks Verlag

Robinson, J. (forthcoming) "Glossary of English lexicographical terms, A-Z" in <u>Workbook on Lexicography: A Course for Dictionary Users</u> by B.A. Kipfer et al. Exeter Linguistic Studies

Zgusta, L. (1971) <u>Manual of Lexicography</u>. The Hague: Mouton