Changing the rules: Why the monolingual learner's dictionary should move away from the native-speaker tradition

Michael Rundell

This paper starts from a recognition that the reference needs of people learning English are not adequately met by existing monolingual learner's dictionaries (MLDs). Either the dictionaries themselves are deficient, or their target users have not yet learned how to use them effectively: whichever view one takes — and the truth is probably somewhere in between — it is difficult to escape the conclusion that the MLD's full potential as a language-learning resource has not yet been realized. This is recognized eg by Béjoint 1981: 219: "Monolingual dictionaries are not used as fully as they should be ... many students are not even aware of the riches that their monolingual dictionaries contain" — a view which has been consistently borne out by any user-research we have conducted at Longman.

There is a variety of responses to this situation. Compilers of MLDs may feel a certain exasperation with the 'ungrateful' students who fail to recognize the very real progress that has been made in adapting conventional dictionaries to their special needs. More positively, a growing awareness on the part of teachers of the importance of developing students' reference skills (eg Béjoint 1981: 220, Rossner 1985: 99 f., Whitcut 1986: 111) is complemented by a clear commitment on the part of dictionary publishers to make their products as accessible and user-friendly as possible. These two developments seem to offer the beguiling prospect of a scenario in which ever more helpful MLDs are ever more skilfully exploited by their users — thus resolving, to everyone's satisfaction, the problem identified at the beginning of this paper. What I want to argue however, is that this rather cosy vision is — for historical reasons — fundamentally flawed. Even the most innovative MLDs are operating within the constraints of a much older lexicographic tradition: their development has been evolutionary, in that an existing model — the Native-Speaker Dictionary (NSD) — has been adapted for use in a language-learning environment. But, it will be argued, the general orientation of this NSD model makes it quite unsuited for this new task. The NSD, despite its pedagogical origins in the pre-Johnson era, has now been settled for well over a century in an essentially 'observationist' role, providing what Dean Trench called (in 1857) an 'inventory' of the language (Read 1986: 43). The MLD — though its perceived role is still very much in a state of evolution — must surely be providing a very different type of service to its users, as a multi-faceted language-learning resource. The view taken here is that it will not begin to do this really effectively until it has emancipated itself from the NSD tradition to
become, not just a user-friendly, ELT-oriented version of the familiar NSD, but a radically different type of book.

What is the NSD tradition?

The dominant concerns of the NSD can be characterized as meaning and coverage. Surveys conducted both by academics and by market-researchers tend to confirm what we would in any case have guessed, that people consult dictionaries principally in order to find the meanings of words.\(^1\) The consensus of dictionary-users is evidently shared by dictionary-makers. Dictionaries are invariably defined — in dictionaries — as being ‘books about meaning’,\(^2\) and this view has — until very recently at least — been taken as axiomatic in most discussions of dictionary writing.\(^3\) If dictionaries are principally concerned with giving word-meanings, it follows that the more meanings a dictionary gives, the better dictionary it is: in other words, the complement of meaning is coverage. One of the founding principles of what eventually became the OED was that “The first requirement of every lexicon is that it should contain every word occurring in the literature of the language it professes to illustrate” (Proposal for the Publication of a New English Dictionary by the Philological Society 1859). A century and a quarter later, with the great work finally complete, its then editor Robert Burchfield was expressing the view that “All levels of dictionary below the OED are amputated versions” (The Bookseller 22 March 1986, 1200). Whatever one thinks of this sort of remark, there is ample evidence, from lexicographic theory and practice to publishers’ promotional literature, to suggest that the provision of the maximum possible number of word-meanings is — whether implicitly or explicitly — the central aim of lexicographers working in the NSD tradition.

How far has this meaning-coverage orientation of NSDs been modified in the design of dictionaries intended for non-native users? It is certainly arguable that, in a gradual process that started with Michael West’s NEW METHOD DICTIONARY (1935), the MLD has departed quite significantly from the NSD model. The specifically learner-oriented features of MLDs are too well known to need much elaboration here, but four areas of real innovation deserve a brief mention: careful control over the language of definition (pioneered by West and a notable

\(^1\) Eg. Greenbaum 1984: “Like the UK students, the US students consulted the dictionary mainly for information on the meanings of words”; cf. Quirk 1974; Tomaszczyk 1979; Béjoint 1981: 215.

\(^2\) Eg. in WEBSTER’S [FIRST] INTERNATIONAL DICTIONARY OF THE ENGLISH LANGUAGE (1894): dictionary — a book containing the words of a language, arranged alphabetically with explanations of their meanings; a vocabulary; a wordbook.

\(^3\) Eg. Zgusta 1971: 21 “It is precisely the meaning of the lexical units upon which the lexicographer concentrates his attention ... Lexical meaning stands in the centre of the lexicographer’s attention”; cf. Kipfer 1984: 43; Burchfield 1986: 18–19.
feature of the Longman MLDs); the provision of information on the grammatical behaviour of words (starting with Hornby's IDIOMATIC AND SYNTACTIC ENGLISH DICTIONARY in 1942 and now a feature of any good MLD); more recently, a greater attention to lexical collocation (for example in LASDE, LDOCE2, and COBUILD); and increasingly, the development of strategies for aiding appropriate word choice, whether through usage notes, synonym sets, or information about pragmatics (for example in LDOCE2 and COBUILD).

These are important and valuable achievements, but the process has perhaps been one of accretion rather than transformation: the traditional NSD model has been augmented but not abandoned in favour of something new. The rest of this paper will aim to show, first, that the basic assumptions of the NSD model have remained to a very large extent intact in current MLDs, and secondly, that this 'resilience' of the NSD has, from a pedagogical point of view, undesirable and unhelpful consequences; finally, the implications of this view for the future design of learner's dictionaries will be briefly considered.

The resilience of the NSD tradition

The influence on MLDs of established NSD methodology is powerfully illustrated by the survival — at least until very recently — of a number of conventions which reflect the diachronic orientation of NSDs and which, so far from being helpful to learners of English, would be regarded by most language teachers as distracting or even obscurantist. By way of illustration, three characteristic features of NSD procedure will be looked at here, the first two briefly and the third in a little more depth.

1. The distinction between polysemes and homonyms: Following historical principles, NSDs treat homonymous words as completely separate headwords (homographs), while dealing with polysemous words in single multi-sense entries, often of very great length. The same convention has been carried over, with little or no modification, to most MLDs, and yet the rationale of this organizing principle is unlikely to be apparent to the end-user — especially when, as occasionally happens, separate homographs exhibit a greater degree of semantic similarity than is found in many polysemous entries. Most learners, for example, would probably see some connection between bay, in its meaning of 'an indentation in the coastline', and bay, when it means 'a recess' (as in a loading bay or a bay window); conversely, few learners would see any connection whatever between the two main meanings of club ('a society that people join' and 'a heavy stick used as a weapon'). Nevertheless, the historically-motivated (but counter-intuitive) organization of the native-speaker tradition has in general been carried over into the MLDs, so that bay appears in LDOCE and ALD as five separate noun
homographs, while *club* appears as just one. Even more confusingly, *drill* ('a tool for making holes') and *drill* ('a form of instruction based on repetition') are grouped together in one homograph, while *drill* ('an agricultural tool for planting seeds') is shown as a separate entry. It is of course highly desirable that multi-sense lexemes should be presented in smaller, more manageable chunks in order to enhance their accessibility. But any strategies for dividing up large dictionary entries must be transparent to the intended user and, in the case of MLDs, should probably be based on some kind of 'semantic flow' idea, rather than being dictated by anything as arbitrary as word history.

2. **The placement of idioms in dictionary entries:** This is a problem that needs little elaboration. The wide differences in idiom-placement policies has been remarked on often enough (eg Stein 1986: 8 ff., Whitcut 1986: 113), and experience seems to show that there is no entirely satisfactory solution to the question of where idioms should be defined (do you define *wear one's heart on one's sleeve* at *wear*, *heart*, or *sleeve*?). There are, however, degrees of 'unsatisfactoriness', and the widely favoured practice of showing idioms at 'the most idiomatic word' (eg LDOCE1, but not LDOCE2) or at 'the most significant word' (eg ALD) or even 'at the word we think the user will probably turn to first' (COBUILD) is difficult to defend when the intended user is a non-native learner. Even more confusingly, the ALD attaches some idiom definitions to the particular senses of a word to which they are supposed to correspond. Thus *be head and shoulders above* is at sense 2 of *head* ('the head used as a measure'), *be unable to make head or tail of* is at sense 3 ('the head on a coin'), and *come to a head* is at sense 18 ('the point of a boil or pimple'). The remaining idioms of *head* are listed at the end of the entry, presumably because they do not correspond closely to any of the lexical meanings of *head* (though in fact many of them do). Essentially this policy requires the dictionary user to already understand the idiom in order to be able to locate it efficiently. Whatever merits this approach may have for the NSD (and the treatment of *head* in ALD roughly follows the COD model), it is surely profoundly unhelpful for non-native users.

The problem, once again, arises from the application to MLDs of diachronically-motivated practices developed in the NSD environment. This is nowhere more apparent than in the area of definition writing, and it is worth looking at this central operation rather more closely.

3. **Approaches to defining:** It would require a separate paper to analyze fully the principles and objectives that underpin/inform the process of definition writing within the NSD tradition; these are in any case far from homogeneous, as Hanks (1979) has shown. Nevertheless it is easy enough to identify certain well-established conventions which are typical of the NSD definition and which have — inappropriately in my view — survived in most MLDs.
(a) *Archaisms in definitions:* The language used in definitions has proved surprisingly resistant to change. For example, the OED definition of the word *hew* is:

to strike or deal blows with a cutting weapon.

This may well reflect the idiom of the 1890s, when the entry was actually compiled, but the use of *strike* and *blow* in these meanings would be regarded by most contemporary English-speakers as archaic — if they appeared anywhere else but in a dictionary (cf. now the advice of Zgusta 1971: 257). What seems to have happened is that a whole range of conventional defining formulae has become 'ossified' in the almost liturgical domain of the dictionary: users accept such formulae, even expect them, in dictionary definitions, even though they would be regarded as stylistically deviant in most other environments. This does not, however, justify unnatural and atypical usage in dictionaries intended for learners. Yet definitions of this kind are very common in MLDs. For example:

**hew** to cut by *striking* or chopping; aim cutting *blows* (*ALD*)
to cut in(to) by *striking blows* with an axe or weapon (*LDOCE1*)

Similar archaisms, like *vessel*, *bear* (verb), *seize*, and *draw* (meaning 'pull') are alive and well in dictionaries of all types.

(b) *The 'idiomatic' definition:* NSDs frequently use defining words in non-central but conventionally understood meanings, for example:

**vulgar** plebeian, coarse in manners, *low* (*COD*).

This rather quaint use of *low*, though reasonably transparent to most native speakers, is a very long way from the core meaning of the word, and might be expected to cause problems if found in an MLD definition. Yet here it is, with others of the same kind:

**base** *low*; dishonourable (*LDOCE1*)
**close shave** *a narrow* escape (*ALD*)
**disreputable** having a bad *name* (*LDOCE1*)
**gossip** *idle*, often ill-natured talk (*ALD*)

(c) *The 'etymological' definition:* This is a definition which is based on a literal translation of a (usually Latin) root word, for example:

**extramural** outside the *boundaries* (of a town) (*ALD*)
**carnal** of the *body* or *flesh* (*ALD*)

The treatment of the word *docile* is a particularly good illustration both of the persistence of this etymological perspective and of the obstacles this can put in the way of understanding how words are actually used. The Latin word *docilis* is defined in Lewis and Short's *LATIN DICTIONARY* (1879) as 'easily taught'. This wording has been preserved to a remarkable degree in English dictionaries (for example: 'teachable' — *COD*; 'easily taught' — *W9*; 'easily trained' — *ALD*; 'easily taught' — *LDOCE1*), yet it fails almost entirely to capture the flavour of
the word as it is used in contemporary English. Compare the following definition of *docile* (taken from LDOCE2):

quiet and easily controlled, managed, or influenced; SUBMISSIVE.

This perhaps shows that learners' needs are better served by definitions based on an analysis of contemporary word use than by what amounts to etymology-through-the-back-door.

(d) *The 'derivational' (or 'truncated') definition:* This definition explains the meaning of word *x* purely in terms of its derivation from word *y,* for example:

- **creative** having the power to create; of creation (ALD)
- **titular** of, belonging to, or related to a title (LDOCE1)
- **extrapolation** is the act or process of extrapolating (COBUILD).

Admittedly, definitions of this uncommunicative type are often supplemented by further semantic information or by contextualized examples. As they stand they can present serious problems for the learner, but it is important to recognize that usability has not, in general, been a central concern of the NSD tradition. Philip Gove—the editor-in-chief of W3—believed that it was better to define *ominous* as 'of or related to an omen' than to attempt anything more discursive and thereby admit into the definition material not present in the base word. Gove represents an austerely purist position (even his own editors did not always follow his advice) but his viewpoint is instructive in that it reveals some of the fundamental assumptions underlying the NSD approach to definition-writing. The twin demands for precision and comprehensiveness are most obviously reflected in definitions like W3's now notorious *door* (cf. Schelbert, in this volume), but their influence has been subtly pervasive in learner's dictionaries too.

(e) *'Contrived polysemy':* Equally baffling to the learner is what has been described by Patrick Hanks as the "reductionist" approach to establishing meaning conventions. This is evident, for example, in the splitting of semantically undifferentiated concepts on the basis of the way these concepts are realized grammatically: thus, an expression like *sing the baby to sleep* is used to generate a separate 'meaning' of *sing* along the lines of 'bring someone into the specified state by singing' (eg W9, CED, LDOCE1). Or again, in the interests of strict substitutability, a word like *shy* is given separate definitions to cover, on the one hand, *a shy person* and, on the other hand *a shy smile* ('showing the quality of shyness').

---

4 Gove 1968: 5; cf. his view that the perfect definition of *export* (noun) is 'that which is exported': "An attempt to improve it by writing 'commodities sold or sent to a foreign country' introduces matter not present in the base verb" Gove 1965: 232.
The inappropriacy of the NSD model

It should now be clear that the NSD model, with its well-established conventions and focuses of interest, continues to exert a pervasive influence on MLDs, determining their content, structure, and general orientation to a very high degree. The question then arises: is this a problem? What is wrong with using the NSD as a model for a more specialized type of dictionary? In the previous section we looked at specific instances of diachronically-based NSD procedures that have carried over into learner's dictionaries, and some of the problems this can cause for dictionary users have already been identified. These can now be summarized as: (1) organizational procedures that inhibit ease of access and sometimes make unreasonable assumptions of the user's knowledge; (2) definition language that presents models of usage which would be unacceptable in almost any other context (eg because obsolescent or excessively formulaic); and (3) entrenched defining techniques which subordinate usability to precision, impede understanding, or convey important semantic information in a needlessly indirect fashion.

Now it is certainly true that contemporary MLDs do not display the Olympian indifference to the user found in some NSDs, and there is every indication that many of the specific problems identified above will gradually be ironed out. But modifications in the way information is presented are not necessarily reflected in changes at the deeper, more philosophical level, where the fundamental assumptions of the NSD remain to a large extent intact.

It is appropriate at this stage to review what these assumptions are. At the beginning of this paper the native-speaker dictionary was characterized as being primarily concerned with meaning and coverage. We are now in a position to refine this characterization, and to contrast it with an alternative description of the nature and objectives of a productively-useful monolingual learner's dictionary.

The primary functions of the NSD are to record and describe. It is a repository of factual data about the various ways a given word is known to have been used ("the dictionary seeks to record what is found to exist" — Introduction to COD, p. vii). It aims at maximum coverage and presents its information in an even-handed, dispassionate way. It is above all word-centred and meaning-centred. That is, it reflects "the notion of words as linguistic units which can easily be isolated and removed from their contexts" (Wikberg 1983: 216). Each word is viewed as a self-contained item capable of delivering a wide range of meanings and freely combinable with other words (within basic grammatical and semantic constraints, of course). Ultimately, the NSD approach is underpinned by a view of language that emphasizes its capacity for innovation, "its ability to make available an infinity of sentences from which the speaker can select appropriate and novel ones as the need arises" (Katz and Fodor 1964: 481). What this means in practice is that the chief value of the NSD is as a reference, a device for decoding language.
The MLD, by contrast, should be seen first of all as a language-learning re­
source, its compilers as much in the business of language teaching as of lexico­
graphy. The MLD's pedagogical role quite properly implies a degree of selection, 
contrivance, and intervention which would be inappropriate in an NSD. While 
James Murray was anxious that the OED should faithfully record every possible 
fact about (for example) a word's pronunciation ("It is a free country, and a 
man may call a vase a vawse, a vahze, a vaze, or a vase, as he pleases" (Murray 
1977: 189)), to apply this approach to a language-learning dictionary would 
amount to a gross dereliction of duty. The MLD cannot simply record "what is 
found to exist": it must select that part of the truth which is judged to be of 
value to its users. It should be clear, therefore, that the model MLD differs cru­
cially from the NSD in that it processes and interprets raw language data, rather 
than merely preserving it for users to refer to. Above all, to fulfil its potential 
as a learning resource, the MLD should abandon the word-centred and meaning­
centred approach of the NSD.

Practical implications for dictionary writing and dictionary design

No doubt it will prove quite difficult to translate these generalizations into speci­
fic strategies, but by way of concluding this paper some practical suggestions will 
be offered as to possible future directions for the MLD.

In the first place it should be recognized that, for the target users of learner's 
dictionaries, meaning is only one — and not necessarily the most important — of a 
wide range of features that go to make up a fully-rounded picture of a word and 
its place in the lexicon. Rather than treating meaning as central, the dictionary 
should give equal weight to all relevant features, including grammar, style and 
register, collocational properties, pragmatic and connotative features, relation­
ships of synonymy and hyponymy, contextual and syntagmatic preferences, and 
so on. In fact, the precise balance and emphasis of a given entry should be deter­
mined by the known (or 'intuited') needs of the learner: for example, it can be 
assumed that words like weep and purchase will pose no real meaning problems 
for learners, but great care will be needed in conveying their rather subtle stylistic 
qualities if the dictionary is to enable learners to use words like this appro­
priately (on weep, see now Jain 1981: 282).

Secondly, learners will be better served by accounts of word-meaning based 
on a 'prototype' approach, which deals in core meanings that admit of minor 
variation and degrees of category membership, rather than by the rigorously 
'criterial' approach favoured by most NSDs (thus sing the baby to sleep would 
not generate a separate category, and in general there would be less 'splitting').

Thirdly — and perhaps most importantly — the MLD must take account of 
the fact that the combinatorial properties of words are in practice very far from 
unlimited. To Katz and Fodor (1964: 481) "the striking fact about language is
the absence of repetition”, and I have argued that this view of words as free-standing and almost infinitely combinable units is implicit in the whole design of NSDs. For the language-learner, by contrast, the really interesting fact about the use of language is the very high degree of repetition that real text actually exhibits — what has been called “the ability of words to predict their own environment” (Jones and Sinclair 1974 passim). In most situations, word choice is quite significantly constrained — for example by strong collocational preferences, questions of style or context, restrictions on subject or object selection or on aspect or number — and for the learner who is attempting to use the dictionary as an aid to encoding, this information is absolutely vital. To give a very simple example, any account in a learner’s dictionary of the word problem should at the very least mention as significant collocates the verbs pose and (especially) solve. Yet, however obvious this may seem, it is a matter of fact that none of the major pre-1987 MLDs (ALD, LDOCE1, and CULD) includes this information. For the MLD, there is much to be learned from the work of Apresyan, Meleč’uk and others on the Explanatory-Combinatorial Dictionaries, while the very large concordanced corpuses now becoming available to lexicographers provide ever more conclusive evidence for a view of language as a system of well-established preferences (see now Hanks, in this volume), in which repetition and predictability are just as important as innovation and creativity. To maximize its generative utility, therefore, the MLD must supply information about the forms and environments in which words tend to appear in most cases: it must, in other words, give an account of what is typical, rather than simply describing what is possible.

This in turn argues for a far more selective approach to the questions of inclusions and coverage. It is becoming increasingly clear that a very high percentage of text is made up of a relatively small section of the total available word-stock. The evidence of the Brown and LOB corpuses, of the AHD frequency study (Carroll et al 1971, esp. pxxviii) and, most conclusively, of the Birmingham corpus, has established that something approaching 90% of sampled text is accounted for by as little as 2% of the word types in a given sample (about 5000 word types, and perhaps as few as 2000 actual lemmas in the case of the 17.8 million word Birmingham corpus). The question then arises whether a productively-useful MLD can realistically be expected to cater for the decoding needs of advanced learners by including vocabulary items of very low frequency, or whether words of this type, and especially those which Zgusta (1971: 68–70) calls ‘autosemantic terms’ (which always mean exactly the same regardless of context), might be dealt with more efficiently by a bilingual dictionary (or perhaps by a separate bilingual component in an MLD?).

The moral of all this is simply that all the characteristics of the traditional NSD should be re-examined by designers of MLDs, and any procedures which do not promote the goals of the MLD — which, as we have seen, are substantially different from the goals of the NSD — should be discarded and replaced. This is not just a question of tinkering with the microstructure: it cannot be assumed
that any NSD feature (even the simple alphabetical headword list with its single-item entries) is necessarily sacrosanct. Such a re-evaluation of dictionary-writing procedures may in the end have quite radical implications for the selection and organization of information about the lexicon, and MLDs of the future may no longer be recognizable as the same species of book as the familiar NSD. None of this will be easy, not least because the conservatism of learner-users — who, as Béjoint (1981) showed, still overwhelmingly use their MLDs for decoding meaning — will have to be overcome if real progress is to be made. But it will certainly be interesting.

References

Cited dictionaries

AMERICAN HERITAGE DICTIONARY (AHD)

CHAMBERS UNIVERSAL LEARNERS DICTIONARY (CULD)

COLLINS ENGLISH DICTIONARY (CED)

COLLINS COBUILD ENGLISH LANGUAGE DICTIONARY (COBUILD)

CONCISE OXFORD DICTIONARY, 7th Edition (COD)

LONGMAN ACTIVE STUDY DICTIONARY OF ENGLISH (LASDE)

LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH (LDOCE1)

LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH, NEW EDITION (LDOCE2)

OXFORD ADVANCED LEARNERS DICTIONARY OF CURRENT ENGLISH, 3rd Edition (ALD)

OXFORD ENGLISH DICTIONARY (OED)

WEBSTERS NINTH NEW COLLEGIATE DICTIONARY (W9)

WEBSTERS THIRD INTERNATIONAL DICTIONARY (W3)

Other literature


Gove, Philip (1968), "On defining adjectives, part 1", in: *American Speech* 43. 5–32.


