Conversational Data and Lexicographic Practice

Barbara Lewandowska-Tomaszczyk

The most frequent source of the material included in dictionaries are written texts. Lexicographers tend to concentrate on words found in written data and to draw their examples from attested written evidence. The aim of the present paper is to try and evaluate spoken data as a possible source of evidence for lexicography.

Is spoken language different from written language, and if so, does this fact matter to lexicographic practice? As is evident from my data, and as also reported in other studies (e.g. Hindle 1983), in conversational English there occur *regular* spoken forms which can be considered ungrammatical or marginally grammatical if we apply the same criteria of grammaticality to the spoken forms as to the written ones. However, such forms are regularly found in spoken materials.

The forms and constructions characteristic of the spoken mode are manifested on several linguistic levels. They may be sentential constructions which do not occur in written language (examples (1)—(4) after Hindle 1983, all other examples quoted in the paper are taken from *A Corpus of English Conversation,* Svartvik and Quirk 1980. The original prosodic markers have been removed; the numbers refer to pages).

(1) What I did I came back and I went to Princeton University, where I had matriculated originally
(2) That's the only thing he does is fight
(3) I imagine there's a lot of them must have had some good reasons not to go there
(4) Oh aaa take the three bus. Will take you to Erie and Torresdale, just walk home. It's not far from there
(5) They just don't apply to Alsatians these epithets but they do to him (725)

The forms found in conversational English may be phrases which are more or less 'unorthodox' in their combinatorial pattern (6) or valence selection (7), or else they may be lexical items whose morphological shape, although the result of a linguistically productive rule, is not quite standard in English (8 and 9):

(6) a very toothless gardener (647)
(7) so I have to smuggle myself in on occasion (472)
(8) English is a very teachy subject (812)
(9) sort of research studenty kind of people (416)

Furthermore, there are also strategies and techniques of sense representation in conversations, which may be worth looking at if one has in mind their possible application in lexicographic practice. This is what I shall attempt in the present paper. More specifically, I shall examine what English language users actually do verbally in order to convey the intended senses of different lexemes (including the 'unorthodox' ones) to their interactional partners and which properties of lexical meanings they consider most salient psychologically and/or socially. I will also try

to show how English speakers extend word senses and forms, and how this material compares with the actual definitions of the elicited senses from some English monolingual dictionaries (COBUILD, CED, C Pbk., LDOCE, ALD, AHD). I will also try to argue that spoken language offers a better way of observing some psychologically more 'real' regularities than does written language. 'Definitions', or, more appropriately, meaning elicitation techniques (cf. Lewandowska-Tomaszczyk 1987), are introduced more spontaneously and tend to display a hierarchy of cognitive salience more clearly than some prescriptively-constructed written definitions.

In most cases a definition in the spoken language is constructed not necessarily in terms of most immediately dominating supertaxa (cf. 'a man is a featherless biped'), but in terms of most *familiar* objects, some of which happen to wear a superordinate label:

(10) this huge Alsatian [. . .] it was really like a sort of beautiful wolf (727)
(11) a wok is one of those big Chinese frying-pans with a rounded bottom you know, that they swizzle things round in (569)

Hesitations and reservations as to the appropriateness of the lexical label to a given referent are more explicit in spoken descriptions, when compared with dictionary definitions, which tend to be either more categorical (i.e. cover only prototypical exemplars), or, on the contrary, give much less definite contextual constraints:

(12) instructing its various academics in how to use, if academics the right word for a Polytechnic teacher — and some would question that I suppose (261)
(12a) LDOCE *academic* 'a member of a college or university, esp. a person whose job is teaching'.

Academic is naturally associated with *academy* and *academies,* which COBUILD defines as

(13) 'some schools or colleges, especially ones that specialize in teaching a particular subject'

and gives as an example *the Soviet Academy of Sciences,* while academies of science, at least in countries such as the USSR or Poland, are research institutions and not teaching establishments.

Meaning elicitations in spoken language employ, like dictionary definitions, typicality features of a referent. However, they maybe more informative on occasion by wrapping up the description in a partly contrastive background (14) and by providing stereotypical associative properties (15):

(14) she was a sort of colourless mouse of a woman but she was very sweet and kind and pleasant and interesting to talk to (99)
(15) all the sailors used to get scurvy, didn't they — which is vitamin C and oranges (598)

Even some apparently paradoxical statements can foreground the intended sense of the item.

(16) S1: they were rather parochial down there
S2: in an international way
S3: very limited, but it's so nice and relaxed there, I mean compared with London, I mean, I found myself going into shops and people smiled at you (569)

The negative evaluation, usually associated with the word *parochial* (cf. COBUILD — 'used to show disapproval') can be mitigated by evoking possible positive attributes of provincial life.

A frequent strategy of defining a lexical sense is by giving its narrower specification as in an example:

(17) John is a very earthy person — he's only interested in beer and sex (592)

LDOCE, which gives a very similar example, seems more explicit in defining the sense of *earthy* than the other dictionaries consulted (cf. COBUILD).

Giving a lexical label to a selected fragment of reality may either precede or follow its description in conversation, whereas in dictionaries the definition conventionally follows the headword. These two strategies, however, are not symmetrical, as is evident from comparing (18) with (19):

(18) S1: I think that's telescoping, I mean
S2: I'm being brief like the man in the conference, I'm trying to cover what I think are the important factors (578)
(19) now you've got a chance there you see, of a group of people with time and with facilities and you could organize them into some sort of coherent body of people with a particular viewpoint — I don't mean a robot viewpoint (477)

An interesting problem with such examples as (19) is that the description preceding the label may correspond to the figurative sense of *robot* (cf. CED, LDOCE etc.). It appears, though, that the definition is not criterial: it is ambiguous in that the speaker of (19) is indeed thinking of a centralized management system, but not to the extent of dehumanizing the people involved. The description given in the example may underlie more than one lexical label, associated with different ideologies and evaluations. (18) is an exhaustive explication of the figurative sense of *telescoping,* which, interestingly, is not included in ALD or LDOCE.

A word or phrase may be elicited in conversation by providing its antonym (20) or by giving an extended contextual comment (21):

(20) S1: never sort of presenting anything sort of nakedly but
S2: always in context (682)
(21) S1: did they just sit silent?
S2: yeah, more or less, or talked to each other in undertones so as not to interrupt the sort of verbiage (667)

The phrase which interests us in (21) is *in undertones.* CED lists this sense of *undertone* but does not give the phrase *in undertones,* while both COBUILD and LDOCE list both the senses and the phrase, though the COBUILD grammatical code (N Count) may be less readily interpretable to a learner than the explicit examples with singular and plural nouns in LDOCE.

A usual method in monolingual dictionaries is the use of synonyms. This technique is also frequently exploited in spoken language:

(22) dunderheads [...] these fools [...] the old plodders [...] old fools (453)

Synonymy, however, is not an 'all or nothing' phenomenon. Rather than looking, unsuccesfully, for absolute synonyms in natural language, we should rather think of degrees of synonymy between pairs of lexical items with the lowest tier firmly anchored in the context (cf. Lewandowska-Tomaszczyk **in press.**). In spoken language and, most notably, in its conversational style, the phenomenon of contextual synonymy, good for a 'practical purpose', is quite frequent:

(23) S1: why did he get bitten with South Africa
S2: yes, why choose South Africa, indeed (734)

get bitten and *choose* are not synonyms in an ordinary sense of the term and they would not be treated as such in an ordinary dictionary. However, there is a clear semantic link between the two, though rather of an antecedent *(get bitten with)* — consequent *(choose)* than of a regular synonymy type. Such sense relations would be easier to encode and account for in a conceptual type of dictionary.

Another sense-relation utilized in meaning explication is the asymmetric entailment relation of the scalar type:

(24) you know Guinness has a slightly sharp taste, a little bit — well it's more than sharp, it's quite bitter (179)
(24a) bitter — [sharp +]
This type of sense relation is employed in most dictionary definitions:
(24b) *bitter*

COBUILD — 'sharp' + 'unpleasant'
C Pbk — 'unpalatable' + 'harsh'
LDOCE — 'sharp' + 'biting'
AHD — 'sharp, acrid' + 'unpleasant'
ALD — 'tasting *like* beer or unsweetened coffee'.

The cross-comparison of the first four definitions with the simile used in ALD leads to the undesirable conclusion that the taste of beer or unsweetened coffee is unpleasant and unpalatable. Such an assertion would be debatable and was most probably not intended by the compilers of ALD. Always of interest to lexicographers is the interface of the scientific universe and reality as perceived by 'the man in the street'. Some such examples occur in the analysed data. One of them is an extended negotiation of the senses of two lexical items *symptom* and *sign* as used by ordinary speakers of English and by doctors of medicine. Here are some passages from the dialogue:

(25) S1 (medical doctor): a symptom is something the patient complains of, the sign is something the doctor elicits [...]
S2 (academic): but I'm talking now about signs and symptoms then and I'm putting them both under the heading symptom
S1 [...] and it is important in medicine to separate them you see (610 [...] 614)

If we analyse dictionary entries for *sign* and *symptom* we find that the potential distinction between these two concepts which gave rise to the medical terminology is not unrecorded in some dictionaries.

(25a) COBUILD

sign 4.1. 'If there is a sign of something, a characteristic feature of it is present, which makes you realize that it is actually happening'

symptom 1. 'A symptom is something wrong with your body or with the way that it works, that is taken as a sign of illness'

Characteristic here is the phrase 'that is taken as a sign' in the definition of *symptom,* which points to the less direct link between a state of affairs and its interpretation than in the case of *sign.* Compare the definition in (25b):

(25b) LDOCE *symptom* 1. 'an outward sign of inner change, new feelings, etc.' 2. 'a change in body or mind which shows disease or disorder',

which does not imply this subtle distinction of senses. In uses other than those indicated above, *sign* is indeed treated as contextually synonymous with *symptom.* It is typically the case that the scientific/technical terminology is a consequence of the elimination of polysemy, i.e. the elimination of a number of senses of a lexical item.

A frequent strategy in meaning exposition in conversations is using an item (often a figurative one) to evoke a 'frame and script' and fill in the details of the frame in the spirit of a parabolic picture (26), a technique not so well developed in lexicographic practice.

(26) He's such a dynamo, he's always driving here and driving there [. . .] he'll burn himself out if he goes at this rate (163)

The technique of frame-building foregrounds the typical properties of an object in a given culture, which are perceived by community members, though not always mentioned in dictionaries; for example, the fact that in our culture a typical church building has pillars suggests the following simile:

(27) a room which curiously has a pillar in the middle of it or several, it looks like a church (830)

Such properties recorded in dictionaries could provide dictionary users with a powerful tool for creative meaning extensions. Such a technique is also indirectly connected with other methods of the *generative* type possible in lexicographic practice such as a 'synaesthesic' type of transfer (the converse of metonymy), where the property of an object is transferred to the agent possessing this property. Cf.:

(28a) COBUILD *patchy* 2. 'it is correct in some parts, but not in all parts, and is therefore considered incomplete and unsatisfactory e.g. *The evidence is a bit patchy, isn't it? [. . .] He has a rather patchy grasp of history'*

(28) persons of foreign origin who speak at conferences are shall we say very patchy in the contents of what they say (586)

Other strategies of a similar type result in

(1) possible extensions (29) and modifications (30) of conventional idiomatic expressions and

(2) exploitation of semantic roles in verbal frames in an unconventional fashion and violating conventional selectional restrictions (31, 32, 33).

(29) our prehistorians in Cambridge were swallowing the whole thing hook line and sinker (370-1)
(30) I'm getting definitely getting very long in the tooth (658)
(31) How to bluff your way through music (625)
(32) It's not that Meryn's totally unreliable he's just generous to a fault (540)
(33) S1: may I borrow your floor (laughs)
S2: if you bring it back (laughs)

I want to conclude this survey of conversational English forms by pointing to other linguistic facts present in my data, which I found lacking in the dictionaries consulted:

a) clipped forms used in colloquial English:
(34) do you want a cig (a ciggie) [cf. cigarette] (656)
b) morphological and lexical variants:
(35) it's difficult to get a word edgewise [cf. edgeways] out of him (750)
c) some lexical and phrasal senses:
(36) cheating *of the first order* (644)
(37) he's a real vicious swine *that number* (729)
(38) *sand blind* (512) [. . .][except for AHD]
(39) *back slapping* sense of humour (659)
d) clearer account of possible lexical ambiguities and polysemy to prevent wrong inferences:
(40) I was rather proud of myself at that stage, very *shiny* uniform — lovely brown (355)
cf. *shiny* CED 1. 'glossy or polished; bright. 2. *(of clothes or material)* worn to a smooth and glossy state, as by continual rubbing.'
cf. COBUILD 1. 'Something that is shiny is bright and looks as if it has been polished; used especially of things with smooth surfaces e.g. [. . .] *He wore a shiny suit'.*
e) clearer marking of conventional evaluative parameters:
(41) if you're white you're called Caucasian — it's *a lot more elegant* than saying white I suppose (666)
f) richer account of relational lexical senses:
good 'average'/'excellent' (606)
much (407), *expensive* (403) etc. [viz. point of reference]
g) coding of some nominal structures:
(42) this great barge of a woman (93)

As I hope this analysis has shown, spoken language is characterised by a more liberal morphology than the written mode, by constituent rather than sentential syntax, and by a partially different organization of semantic information in sense-explanations. I believe that a truly generative dictionary must take into consideration this linguistic material which is *regularly* attested in spoken language, even if it is rarely used in its written counterpart.

References

Cited dictionaries

THE AMERICAN HERITAGE DICTIONARY OF THE ENGLISH LANGUAGE (AHD). 1969. William Morris and M. Berube. Boston: Houghton Mifflin.

COLLINS COBUILD ENGLISH LANGUAGE DICTIONARY (COBUILD). 1987. John Sinclair et al. (eds.). London and Glasgow: Collins.

THE COLLINS ENGLISH DICTIONARY (CED). (Collins Dictionary of the English Language, Second Edition). 1986. Patrick Hanks et al. (eds.). London and Glasgow: Collins.

THE COLLINS PAPERBACK ENGLISH DICTIONARY (C Pbk.). 1986

LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH (LDOCE). 1978. Paul Procter et al. (eds.). Harlow & London: Longman.

OXFORD ADVANCED LEARNER'S DICTIONARY OF CURRENT ENGLISH (ALD). 1974. A.S. Hornby et al. Oxford: O.U.P.

Other Literature

Hindle, D. 1983. 'Deterministic parsing of syntactic non-fluencies' in *Proc 21st ACL Meeting*, Cambridge, Mass. 123—128.

Lewandowska-Tomaszczyk, B. 1987. *Conceptual Analysis, Linguistic Meaning, and Verbal Interaction*. Lódz: Acta Universitatis Lodziensis.

Lewandowska-Tomaszczyk, B. in press. 'Meaning, Synonymy, and the Dictionary' in J. Tomaszczyk and B. Lewandowska-Tomaszczyk (eds.). *Meaning and Lexicography*, Amsterdam: Benjamins.

Svartvik, J. and R. Quirk. 1980. *A Corpus of English Conversation*. Lund: CWK Gleerup.