How to choose stimuli
for experiments on lexical ambiguity?
A comparison of data sources for psycholinguistic experiments

Daniela Marzo & Birgit Umbreit

University of Tübingen, SFB 441

daniela.marzo@uni-tuebingen.de

birgit.umbreit@uni-tuebingen.de

In psycholinguistic experiments on lexical ambiguity (e.g. priming, eye-tracking, rating and sorting tasks) researchers are in need of suitable polysemous stimuli. For example, it is important to assure that the subjects are familiar with the word meanings that are investigated, because unknown stimuli distort the results. Thus, the most reliable materials for such experiments are relatively salient meanings that are easily accessible for the informants. However, in contrast to what one might expect it seems that this issue is often neglected: some authors do not even give the sources of their materials (e.g. Williams, 1992, Brisard, Van Rillaer, and Sandra, 2001). Still, different ways to choose stimuli can be observed in the literature:

(a) by introspection

(b) by corpus analysis

(c) by consulting dictionaries

(d) by production tasks

Though introspection is one of the most common methods, it is not without problems. It is generally known that the linguistic behaviour of native speakers and their introspective judgements of this behaviour do not necessarily coincide. For example, speakers can be subject to unconscious normative tendencies or other disturbing factors. Therefore the introspective judgements of a single linguist, which are not necessarily consensual, can certainly not be representative for the whole speech community (Bergenholtz and Mugdan, 1990:1614) and consequently cannot lead to reliable data.

A slightly more objective method is corpus analysis. The advantage with respect to introspection is that it is normally based on a large number of sources representing different text types. The linguist no longer makes up the meanings by himself, but can rely on more objective data. Still, there is a problem with these data: often the textual context or the communicative situation are too vague to allow for a clear interpretation of the words one is looking for (Raukko, 2003:165). This has the effect that word meanings either remain ambiguous or have to be defined in a very general

way, which, because of its abstractness, perhaps does not correspond to the mental representation of the typical speaker.

The third method, the consultation of dictionaries, provides even less appropriate stimuli than the other two. Problems start with the selection of the dictionary from which the meanings are taken. In view of the large variety of monolingual dictionaries in certain languages, it is extremely difficult to choose one of them, as the principles on which their meaning differentiation relies diverge considerably. Consequently, they do not necessarily judge the same meanings as central (Gorcy, 1989, Werner, 1989, Schlaefer, 2002:100-102). Moreover, after having chosen a dictionary, first there still is the problem of deciding which of the numerous meanings to use for the experiments and which to omit and second, the meanings they contain do not necessarily correspond to those the informants are aware of.

So far, the most objective method to determine meanings of polysemous words has been production tasks in which informants report the meanings that come to their mind either by formulating disambiguating sentences (e.g. Raukko, 2003) or by defining the meanings directly (e.g. Durkin and Manning, 1989). The advantage of both procedures is that one can be sure to get the meanings that are most salient for the speakers. Besides, if they are conducted under the same conditions, they are also suitable as a basis for cross-linguistic studies as they guarantee the comparability of the material. Nevertheless, our own studies have shown that each method on its own can sometimes lead to ambiguous results. The supposedly disambiguating sentences frequently permit different readings, while the main problem with the definition task is that it is often too hard to carry out for untrained subjects (Dunbar, 2001:2-3).

After illustrating our criticism of the way stimuli are usually selected, we will present in this talk a refined production experiment, the so-called Sentence Generation & Definition Task, which combines both of these methods: informants do not only have to formulate disambiguating sentences, but are also asked to add a definition or paraphrase of the meaning they have in mind. This combination has turned out to be extremely advantageous in the course of our experiments on lexical ambiguity for several reasons: on the one hand, it elicits objective and easily interpretable data that can be used as a reliable basis for all sorts of linguistic experiments using ambiguous words. On the other hand, the risk of presenting our informants with materials they do not know is considerably reduced when we choose our stimuli from the data collected with the help of the Sentence Generation & Definition Task. We will support these claims by presenting a comparison of the results we obtained by applying the methods (a) – (d) to a set of Italian high-frequency words.

References

Bergenholtz, H. and J. Mugdan (1990). Formen und Probleme der Datenerhebung II: Gegenwartsbezogene synchronische Wörterbücher. In F.-J. Hausmann et al., eds.,

Wörterbücher/Dictionaries. Ein internationales Handbuch zur Lexikographie/An International Handbook on Lexicography, vol. 2, 1990, pp. 1611-1625.

Brisard, F., Van Rillaer, G., and D. Sandra, (2001). Processing polysemous, homonymous and vague adjectives. In H. Cuyckens and B. Zawada, eds., Polysemy in Cognitive Linguistics 2001, pp. 261-284.

Dunbar, G. (2001). Towards a cognitive analysis of polysemy, ambiguity, and vagueness. Cognitive Linguistics, 12/1: 1-14.

Durkin, K. and J. Manning (1989). Polysemy and the subjective lexicon: semantic relatedness and the salience of intraword senses. Journal of Psycholinguistic Research, 18/6: 577-612.

Gorcy, G. (1989). Différenciation des significations dans le dictionnaire monolingue: problèmes et méthodes. In F.-J. Hausmann, et al., eds., Wörterbücher/Dictionaries. Ein internationales Handbuch zur Lexikographie/An International Handbook on Lexicography, vol. 1, 1989, pp. 905-917.

Raukko, J. (2003). Polysemy as flexible meaning: experiments with English *get* and Finnish *pitää.* In B. Nerlich et al., eds., Polysemy. Flexible Patterns of Meaning in Mind an Language 2003, pp. 161-194.

Schlaefer, M. (2002). Lexikologie und Lexikographie. Eine Einführung am Beispiel deutscher Wörterbücher, Erich Schmidt Verlag, Berlin.

Werner, R.(1989). Probleme der Anordnung der Definitionen im allgemeinen einsprachigen Wörterbuch. In F.-J. Hausmann et al., eds., Wörterbücher/Dictionaries. Ein internationales Handbuch zur Lexikographie/An International Handbook on Lexicography, vol. 1, 1989, pp. 117-930.

Williams, J. N. (1992). Processing polysemous words in context: Evidence for interrelated meanings. Journal of Psycholinguistic Research, 21/3: 193-218.