Position paper for Tbilisi Round Table

Patrick Hanks

Introduction

In this brief contribution, I suggest that lexicography has -- or should have -- a central role to play in linguistics and other language-related academic disciplines, including philosophy of language, computational linguistics, foreign-language teaching, translation studies, and stylistics. There are, of course, many different kinds of dictionary, but my remarks apply, with certain differences of emphasis, to fundamental principles affecting all the main lexicographical genres -- notably synchronic dictionaries, learner's dictionaries, dictionaries of historical record, and bilingual dictionaries.

Insofar as lexicography does not currently have such a role -- and there are very few places in the world where it does -- the causes may be attributed partly to the narrowness of focus that is characteristic of academic research in the humanities, and partly to false beliefs perpetuated by dictionaries themselves. The increasing availability of empirical evidence -- corpus evidence -- calls traditional lexicographic practices into question. For this reason, in order to have any chance of fulfilling this role, lexicography must question received practices and adopt new approaches, especially with regard to the writing of definitions, the representation of collocational preferences, and the association between meaning and phraseology.

It is not necessary here to make a distinction between lexicography and lexicology. Every aspect of the empirical study of words has a potential contribution to make to the understanding of language and should play a central role in teaching and research. But for this to become possible, a change of direction in needed, both in the academic community and in lexicography. To some extent, this is already happening in dictionaries. Lexicography has already joined forces with corpus linguistics, but it must also learn lessons from theoretical research. It must apply the lessons of prototype theory (Rosch), stereotype theory (Putnam), the idiom principle (Sinclair), frame semantics (Fillmore), construction grammar (Fillmore, Goldberg), and others. We must take seriously Wittgenstein's dictum, "Do not look for the meaning; look for the use". Above all, we must abandon the Leibnizian goal of stating necessary and sufficient conditions for each meaning of each word, a goal that still pervades definition writing, even though it has been shown to be incompatible with accurate language description. We must allow ourselves to be forced by the nature of language itself to adopt a fuzzy model of word meaning, in which definitions or translations are probabilistic statements of stereotypical meanings determined by particular phraseological contexts. Such contexts must be stated -- or rather typified -- explicitly. The relationship between the core meaning of a word and its typical contexts is overdue for systematic re-examination, word by word, language by language, in the light of corpus evidence.

Inadequacy of linguistics

Academic research in the humanities typically aims at depth: i.e., the aim is to say everything there is to say about the chosen topic, however narrow it may be. By contrast, lexicographers normally aim for breadth of coverage. Typically, they feel obliged to say something (but, usually, not very much) about every word -- even if the word is exceedingly rare or if they don't know what to say. It is no longer permissible for lexicographers to say, as Johnson (1755) did when he stumbled across the word *minnock* in Shakespeare (*Midsummer Night's Dream*, III.ii.19), "Of this word I know not the precise meaning."

The vast number of words in a language and the infinte number of possible meanings that they can be used to generate are forces that have combined to convince some linguists that nothing useful can be said about words and meaning and that dictionaries are therefore necessarily superficial or trivial. This impression has been reinforced (in the English-speaking world at least) by the great attention paid by dictionary marketing people to rare and unusual words. For example, logging into the Merriam-Webster on-line web site on 14 August 2016, I am offered a "word of the day", namely *tog,* defined as "to dress especially in fine clothing -- usually used with *up* or *out*", with an example "her teenage son … was togged out in a tuxedo." The word is not in my active vocabulary, and I do not miss it, though it is in many dictionaries. There are many thousands of such rare and unusual words recorded in dictionaries (and dutifully copied into WordNet). I am not saying that there is anything harmful in the wider dissemination of such oddities, but rather that, if lexicography is going to be able to make an acceptable contribution to the study of language, more serious matters must be addressed as well.

Noam Chomsky (surely the greatest living polemicist) famously dismissed lexicography as "mere butterfly collecting", and one can see how he could have got this impression. In 1972 he dismissed Saussurean structuralism as an "impoverished and thoroughly inadequate conception of language." Against this, anyone interested in meaning might be forgiven for dismissing Chomskyan linguistics, which has so little that is useful or accurate to say about meaning, as an even more impoverished and inadequate conception of language.

A result of such polemics is that, in many universities, courses in linguistics do not pay adequate attention to words, phraseology, and meaning. Lexicography can help to change this situation, but it will take time, perseverance, and, I suggest, a continuing change of focus.

Inadequacy of traditional dictionaries

Traditional English dictionaries typically present misleading information about the meaning and use of words. The main problem here is that dictionaries do not say enough about the phraseological preferences (past or present) of each word, and so give a false impression that words combine freely, governed only by the rules of grammar. The evidence from corpus linguistics suggests that many words do not do this. They have strong collocational preferences, which are still, alas, largely ignored by dictionaries. As a result, dictionaries tend to systematically underrestrict the definiendum. I will give just one example. In most monolingual English dictionaries, the verb *put* is given up to two dozen senses, plus another

dozen phrasal verbs and idioms. However, there are no systematic attempts to explain how one sense can be distinguished from another. The definitions are written as if they could contribute freely in all possible syntactic combinations. Thus, one of the definitions of **put** in *Collins English Dictionary* (sense 15) is "to throw or cast". This seemingly corresponds to Merriam-Webster's sense 1 c (2), "to throw with an overhand pushing motion" However, the word does not normally have this meaning. As far as I have been able to ascertain, **put** normally has this meaning only in the expression *put the shot* (referring to an event in field sports). At least MW has the goodness to give *put the shot* as an example. Collins leaves the reader utterly baffled.

Similarly, Merriam's definition 2a of this verb, "to cause to endure or suffer something" is an equally severe underrestriction. The word only -- or perhaps we should say, normally -- has this meaning in the phrase *put someone to death*. You cannot, in English, *put someone to a prison sentence or *put someone to pain.

Consider the user

Young people coming fresh from high school to university and opting to study something called "Linguistics" might reasonably expect to be taught something about how words are used to make meanings.

Ordinary people picking up a dictionary to find out about the meaning or correct use of an unfamiliar word and finding, as so often happens, that the word has several meanings, are entitled to ask how they are supposed to distinguish one meaning from another. (Standard English dictionaries don't tell them.) The situation may be different in other languages, e.g. Russian, if the work of Mel'čuk and Apresjan has had any effect. It should also be noted that, thanks to the efforts of people like Sue Atkins, the best bilingual lexicography already pays far more attention to phraseology than its monolingual cousin.

What is to be done?

In my view, the remedy is theoretically simple but implementing it faces immense practical difficult. The study of lexis, including the relationships between word meaning and phraseological meaning, must aspire to a central role in university courses, augmenting or replacing that of syntactocentric theorizing. Syntax must be put in its place, as the glue that holds words together when they are used to make statements and questions about what happened, or who did what to whom. Methods based on speculation and introspection must be replaced by empirical analysis of corpus data. Methods aiming at certainty must be replaced by methods aiming at a high degree of probability.

For this to become a practical reality, lexicography must first set its own house in order, by investigating in depth, word by word, the relationship between core literal meaning and phraseology and by being prepared to reconsider the almost universal practice of defining by substitutable synonymy. Then, research councils and finding agencies must be persuaded to provide money for new, practical lexicographic approaches to understanding and recording how people use words to make meanings. Finally, symbiosis must be sought with sympathetic souls in positions of authority in university departments teaching linguistics and language-related subjects. This will take skill, and decades-long perseverance.