

## **Linguistic Corpora in Philosophical Analyses**

*Call for papers: Société de philosophie analytique (SOPHA)  
SOPHA congress 2012, 4–6 May 2012, Paris*

Ordinary Language Philosophy has largely fallen out of favour and with it the belief in the primary importance of analyses of ordinary language for philosophical purposes. Still, in their various endeavours philosophers not only from analytic but also from other backgrounds make reference to the use and/or meaning of terms of interest in ordinary parlance. In doing so, they most commonly appeal to their own linguistic intuitions (in the sense of their active knowledge of the object language) with respect to sayings, figures of speech, idioms, collocations and syntactic constructions. This is problematic in two ways. Firstly, the one who thinks of the examples usually has an investment in specific hypotheses that might bias his examples. Secondly, everyone's linguistic competence is limited and can therefore not serve well as a primary data base. This is, of course, one critique that has been levelled against the historical Oxford-based Ordinary Language Philosophy. It has been pointed out quite rightly that its proponents have occasionally failed to notice that their sophisticated use of English is not to be confused with the use of English as a whole.

There are a number of ways in which philosophers could and do overcome the limits of individual linguistic competence.

Quite commonly, the appeal to individual intuitions is simply supplemented by reference to a dictionary or a number of dictionaries, especially when diachronic variations, but also when synchronic meaning in the meaning of a concept of interest are of any importance. However, dictionaries are not to be trusted unquestioningly. First of all, they are, of course, not without mistakes or oversights. New dictionaries also partly rely on older dictionaries, thereby preserving not only past knowledge but also past mistakes. Not all dictionaries record recent developments (the most commonly used printed editions of the large dictionaries are dated). And, finally, dictionary definitions are not intended for the same purposes as are the products of conceptual analyses or explications.

Thus, if ordinary language is important with respect to some philosophical endeavour, there is the need for some basis on which the dictionaries' definitions can be tested, corrected and extended. Put another way, we need an independent basis for formulating hypotheses about the use of concepts in question, and we need independent means to test these hypotheses.

In recent times, it has become quite popular to point to the results of rough-and-ready internet search engine queries with respect to the question whether certain uses are possible. Apparently, the attempt is to surpass the limits of one's own linguistic intuitions by appeal to the factual uses of language. This attempt is recommendable if conceptual analyses are to be based on ordinary language, but its execution is wanting. Most importantly, internet search engines offer a limited number of search algorithms. With respect to word combinations, these algorithms are radically limited. The internet is, of course, not linguistically annotated, and thus lacks information that could otherwise be employed for sophisticated hypothesis

testing. Furthermore, English is used on the internet by a large number of speakers who have limited linguistic competence. The fact that, e.g., certain constructions are used is thus not in itself good evidence for their acceptability. And, lastly, internet search engines cannot offer any statistical information that is precise enough to be useful.

Another option – recently put to sophisticated use by experimental philosophers – is to ask informants, i.e. optimally, to use a questionnaire in a controlled setting to obtain informants' views on how a specific concept is used, or on whether certain constructions are or are not objectionable, or on the meaning of a specific concept. But even if the problem of the limits of linguistic competence could be addressed in this way (and I am not certain that it can be addressed satisfactorily in this way), a major flaw of this method is to draw the informants' attention to their use of the language and to thereby invite answers that do not provide information on how informants use a specific concept but on how they believe they should or do use the concept.

All of these problems can, as I want to argue, be avoided by employing computer-based linguistic text corpora. Nowadays, there is a variety of such corpora that can be accessed easily online. The search interfaces are for the most part intuitive and usually offer a variety of partly very powerful search algorithms. With regard to philosophical conceptual analyses, linguistic corpora can be put to a number of uses. They provide data on the basis of which hypotheses can be formulated; they provide data to confirm or falsify hypotheses; and they can serve as a collection of data that can be used to exemplify or illustrate specific usages of interest (cf. Bergenholtz & Schaefer 1985, 281). All corpora provide some of the linguistic context (one or two sentences) in which the queried strings have been found, quite frequently a wider context is provided. This allows for in-depth analyses of interesting specific uses of concepts. It also provides rich insights into the variety of situative contexts in which the phenomenon referred to by the concept occurs. Additionally, such contexts often provide excellent raw material for thought-experiments with regard to the concept or the phenomenon in question.

Thus, instead or at least in addition to appealing to dictionaries, internet queries or questionnaires, philosophers should employ computer-based linguistic corpora, at least if the importance of ordinary language in respect to a specific philosophical endeavour is stressed.