Stefanie WULFF, *Rethinking Idiomaticity. A Usage-based Approach*

Antoine Consigny
REFERENCES

Stefanie Wulff

1 The book is 240 pages long, including appendices (56 pages), a 9-page list of references and two indexes (author and subject), and is the product of a PhD thesis. As such, it is suitable mainly for researchers working in or familiar with the field of corpus linguistics, cognitive linguistics or both. Though the methods used are explained in sufficient detail, some knowledge of statistics and quantitative corpus linguistics is useful to fully grasp the implications and relevance of the findings.

2 After a brief introduction, which outlines the theoretical framework (usage-based, cognitive linguistics) and the research questions (idiomaticity and its relation to semantic and other types of variation), 6 chapters present the various stages of the study – theoretical and methodological issues (1 and 2), the question of compositionality (3), then the bulk of the study appears in Chapter 4, on flexibility measures, 5 and 6 being further discussion of the results. A seventh chapter concludes the study.

3 The first chapter is in two parts. It starts with a review of the literature on idioms and idiomaticity. Wulff rejects the generative dichotomy between opaque (idiomatic) and non-opaque (non idiomatic), preferring a phraseology-oriented, cognitive-linguistic view that it is a matter of degrees rather than of absolute belonging or non-belonging: ‘any multi-word expression can be placed on a collocation-idiom continuum according to its
In the second part of the chapter, she presents and explains the approach taken in the book – a constructionist approach (following Goldberg, 2006), which is at the same time cognitive and usage-based. In this approach, idioms become the central part of grammatical study rather than being relegated to a peripheral role as was the case in early generative studies. They are defined by their frequency of use and their recognition as units by speakers; opacity becomes a sufficient but not necessary criterion.

The second chapter describes the overall method followed, as well as the data used in the study.

First, the question of corpus linguistics is addressed, this being the field of this study, and more specifically the question of quantitative corpus linguistics (Stefanowitsch and Gries, 2003, 2005 etc.). This approach is described as ‘a methodology that allows for the investigation of theoretically informed hypotheses at a level of sophistication that is not possible by employing traditional corpus linguistic methodology’ (p. 21). It relies first and foremost on real language data, and in particular on frequency of use, since that is seen as a major factor in the recognition of a sequence as a construction: frequency of use induces frequency of exposure. This also justifies the link between corpus linguistics and cognitive linguistics which is claimed in this section.

Secondly, she presents the data gathered for the study. They consist of a sample of phrases (V NP-constructions, i.e. containing a verb and a noun phrase, e.g. cross X mind, meet X eye, etc.) included in the Collins COBUILD Dictionary of Idioms (1995), to which are added a number of frequent phrases not included (e.g. call X police). Informants are asked to judge the idiomaticity of each phrase, and the results are then normalised for easier comparison.

In Chapter 3, Wulff looks at the question of compositionality. After looking at previous approaches, she adopts Berry-Rogghe’s view (1974) that compositionality can be measured by comparing the phrase’s collocates with those of the constituent parts. She justifies this view by pointing out that if the phrase and part(s) have many collocates in common, then the construction is highly compositional. Berry-Rogghe’s measure of the contribution of a particle to the semantics of a verb-particle construction is adapted and extended to take account of both parts. She concludes that compositionality is a matter of degrees, and that it is not systematically linked to one type of constituent part or another.

The next chapter is devoted to the question of the flexibility of V NP-constructions. Drawing on previous work (Barkema 1994), which she extends to suit her own needs, Wulff conducts a very thorough and detailed study of the different variation parameters – syntactic, morphological and lexical, both of the verbal part and the nominal. In other words, the scope goes well beyond usual parameters (a typical example being Fraser 1970) such as the possible use of the passive, although that is also included. The results for each parameter are presented in tables that give a picture of the degrees of flexibility of the different phrases. They include two kinds of comparison – first, with a ‘baseline’, i.e. the mean behaviour of all V NP-constructions outside the data sample, and also a comparison with compositionality (Chapter 3). She concludes that flexibility varies between phrases, and also between parameters for a given phrase. On the whole it is lower than the baseline, especially for the passive voice, thus confirming findings of previous idiomaticity research, but no systematic correlation is found with (non-)compositionality.
Chapter 5 presents an attempt to find which of the different parameters are most prominent. In order to do so, Wulff looks for clusters of parameters, which she calls Principal Components (PC). The method is described as follows: the Principal Component Analysis is a statistical procedure by which variables are grouped together when they correlate highly with each other, i.e. when they show commonalities. The aim is to identify clusters to reduce the overall number of variables, but not the informational yield. From the 20 parameters under consideration, she identifies 8 PC, accounting for almost 80% of the variance from the expected/normal patterns (the baseline). She then concludes with a discussion on which parameters are most important.

The final chapter compares the (usage-based) variation parameters of Chapter 4 with (intuition-based) judgements on idiomaticity as described in Chapter 2, the aim being to find out whether the two correlate. The analysis is done in two steps. First, the parameters taken together are compared to idiomaticity judgements with a multiple regression analysis. They are shown to correlate highly, and to account for ‘nearly 80% of the variance in the average idiomaticity judgements’ (p. 158). Secondly, the relative weight of each separate parameter in the overall correlation is looked at. This shows that some parameters seem to be far more important than others. 9 parameters are thus identified, 8 of which also form the most important PC (Chapter 5). This is taken by Wulff to mean that speakers rely more on them, and to confirm her assumption that usage-based, quantitative corpus linguistics can shed light on cognitive, construction grammar, even though the results may not fit one’s intuition. She concedes that ‘the corpus-linguistic study results do not provide a definite answer [to the question which parameters speakers rely on], but still considers that ‘they suggest different possibilities’ (p.160), which she then turns to in the remainder of the chapter.

As a conclusion, Chapter 7 sums up the results. This usage-based study goes beyond previous studies, especially since many variation parameters are considered concurrently, and shown to contribute, albeit to varying degrees, to idiomaticity. Wulff then presents a new, multidimensional model for the representation of constructions, taking into account all aspects of a construction from schematization to idiomaticity.

Throughout the book, S. Wulff uses fine-grained analyses and very precise statistics, giving her results weight both from a corpus linguistics perspective and from the cognitive linguistics perspective. The link between the two, made particularly explicit in Chapter 2, is however not universally accepted, nor always obvious in the study itself. For instance, Teubert (2005) considers that ‘corpus linguistics and cognitive linguistics are two complementary but ultimately irreconcilable paradigms’; but the two views of corpus linguistics are different. To Teubert, it is, as Wulff herself puts it (p.21), ‘a real paradigm’ whereas in her view it is ‘a methodology’ used in cognitive linguistics. Yet, one point which she develops can be taken as quite close to Teubert’s approach rather than a cognitive linguist’s approach – the question of the semantics of constructions and their compositionality (Chapter 3). She defines the meaning of a unit as the collocates it has. If it is generally accepted that collocates (among other things) enable one to distinguish between different meanings of a unit, it does not necessarily follow that each meaning is defined by its collocates. For example, if a meaning is rare, it may be difficult to find its collocates, even in the BNC, which is used in Wulff’s study. Does it follow that a construction that includes this particular meaning is necessarily non-compositional? The question is also relevant (and indeed mentioned by Wulff) for very frequent and/or very polysemous words, in particular verbs, like take. The answer she
gives may be taken as possible, but there lacks conclusive evidence on this question, and there is still much room for debate. Similarly, though the difference between non-compositionality and semantic opacity is mentioned, the latter is more or less left out of the discussion; yet, idiomaticity judgements given by ‘naïve native speakers’ (p.32) may well be influenced by the commonly held idea that idioms are semantically opaque (though it has been shown that they need not be, which Wulff accepts). The idiomaticity judgements are compared with statistically defined and justified parameters, but unlike the latter, the former are not justified in any way, leaving aside the possibility that opacity was important for the informants. As mentioned earlier, she does not claim to give ‘a definite answer’. However, the link with previous cognitive linguistics research which she does claim is less obvious here, because the cognitive component is linked (somewhat artificially, it seems to me) to pure quantitative corpus linguistic results, and part of it is even left out. This being said, the book is a considerable piece of research, whose results are important and constitute a real breakthrough in linguistics. The theoretical discussion is far from over. Rather than providing definitive answers, this work should be seen as providing food for thought, an excellent intellectual stimulant.

BIBLIOGRAPHY


AUTHORS

ANTOINE CONSIGNY

Antoine Consigny est maître de conférences au département d’Études Anglaises et Nord-Américaines de l’Université de Strasbourg, où il enseigne la linguistique anglaise de la 1ère année à l’agrégation. Il a soutenu à Liverpool en 2002 une thèse sur la sémantique des verbes à particules, et a poursuivi ses recherches en sémantique, en particulier en sémantique verbale, et en grammaire.