

**Birgit Steinbügl.** *Deutsch-englische Kollokationen: Erfassung in zweisprachigen Wörterbüchern und Grenzen der korpusbasierten Analyse.* 2005, VI + 189 pp. ISBN 3-484-39126-X. ISSN 0175-9264. Lexicographica. Series Maior 126. Tübingen: Max Niemeyer. Price: €98.

This study investigates the role of collocations in dictionary use, and the extent to which users' needs are taken into account in the process of dictionary writing. Steinbügl decided to concentrate on bilingual dictionaries, because this type of dictionary is relatively less well explored in metalexicographical literature. German-English examples are analysed and evaluated. Instead of selecting examples randomly, she uses a comparative corpus of 200 collocations she put together herself in accordance with scientific reasons explained in detail. She questions the selection of collocations from existing corpora for her purposes, because these corpora are based on competing collocational theories. In order to come to meaningful conclusions, she prefers to delineate her own research approach, also however investigating the structures of bilingual dictionaries and dictionary articles, as well as situations of dictionary use.

In the first chapter of her book, Steinbügl gives a theoretical exposition by discussing different views on the notion of "collocation" (2005: 3). She firstly explains Hausmann and others' approach, which she describes as a "lexicographically-didactically oriented theory" (2005: 12). Hausmann distinguishes between non-fixed and fixed combinations of words and figures of speech. A fixed combination is encountered when its meaning cannot be explained by simply considering the semantic values of the individual components; this means that fixed expressions are to be interpreted as one sign. For example, the expression *chambre forte* is not automatically understandable for non-native speakers of French, because they cannot deduce the meaning from the individual components in the expression. Knowing the meaning of the individual words does not help them to interpret the fixed expression, which applies to the receptive as well as the productive use of a foreign language. For dictionary users, it is sometimes difficult to locate the explanations of fixed expressions, especially when dictionaries differ in their methods of placing these expressions.

Hausmann (1984: 399) describes three types of non-fixed combinations. Firstly, there are free combinations or co-creations, which can be produced at any time, and in which each individual component has an independent meaning. Secondly, he identifies collocations or so-called "affine combinations". He regards "affinity" as the "tendency of two words to appear in combination" (1984: 398). Collocations are not creatively formed; rather, they are "pieces" memorised by the speaker because they so often occur in combination. As the combinatory possibilities of the collocation partners are restricted, the use of the words are considerably reduced. It is therefore not possible to adequately explain the meanings of these combinations by semantic means. Thirdly, there are so-called "contra-creations" which occur only rarely, e.g. when novelists

deliberately break semantic and/or grammatical rules in order to combine words normally not combinable, usually for metaphorical effect.

Steinbügl draws strongly on Hausmann's view, especially using his second category as her point of departure. She emphasises (2005: 6, note 6) that the components of collocations need not be juxtaposed together in texts; Hausmann (1985: 124ff) accordingly uses the expression "collocation span" to indicate that words appearing relatively far away from each other in a sentence could still form a collocation. He also differentiates between "basis" and "collocator". Steinbügl explains this difference with *guilty conscience* as example. The noun *conscience* can be understood as the basis, and the adjective *guilty* as the collocator which completes the collocation. The basis which is semantically autonomous, can be defined without a context; the collocator can only be defined in terms of the basis.

From another viewpoint, the so-called "statistically-oriented approach", propagated by linguists such as John Sinclair and Göran Kjellmer, the presence of collocations is indicated when two linguistic entities appear together in the same text. When this simultaneous occurrence is found more often than expected, it is called a "significant collocation". This statistical approach is closely linked to computer-assisted corpus linguistics. In her study, Steinbügl also analyses corpus-based collocations, even though she believes that frequency alone cannot be taken as criterion for the identification of collocations. She proceeds from the assumption that the intuition of the linguist or lexicographer cannot be replaced by machine-based investigations, even though data obtained by computers can be more reliable than the intuition of a single mother-tongue speaker.

In British contextualism, Firth claims that the situational context in which an utterance is made, is of crucial importance to the meaning of language. Apart from this extra-linguistic dimension, meaning also has a pure linguistic level, on which collocations play a decisive role. He uses the notions of "habituality" with which words occur together, and the "mutual expectancy" of words to appear together. Differing from Hausmann's view that a hierarchical relation exists between basis and collocator, he believes that the meaning of both partners in the combination is equally influenced by the collocation.

The "text-linguistic" notion of collocation represented by scholars such as Halliday and Hasan uses the concept "collocational cohesion", which can be exemplified by *hair, comb, curl* and *wave*, which, as extra-linguistic entities, would form a combination and therefore a collocational cohesion. According to Steinbügl (2005: 15), however, the fact that no distinction is made between linguistic and extra-linguistic relations is problematic.

One of the problems facing lexicographers of bilingual dictionaries is that different languages have different ways of forming collocations. This also makes it extremely difficult for foreign language learners. Another aim of Steinbügl in this study is not only to look at pairs of two words, but also to investigate idiomatic and fixed expressions.

After the initial exposition of her theoretical considerations, Steinbügl discusses the theoretical foundations of bilingual dictionaries in the second chapter. She gives brief overviews of structural aspects such as the macrostructures of bilingual dictionaries and the microstructural components in bilingual dictionary articles. She also considers problems which may arise in the writing and use of dictionaries. She draws examples for her analyses from the following bilingual dictionaries: Duden Oxford bilingual dictionary (DOG-E2), Langenscheidts' bilingual dictionary with German and English (LHW-E 2000), Langenscheidts' bilingual dictionary with German and French (LHW-F 2000), Pons Collins dictionary with German and English (PCGW-E4) and Pons bilingual dictionary with German and French (PGW-F1). Later in her study, she then uses features of these examples for quantitative analysis.

Steinbügl analyses the niche-alphabetical macrostructures from some of these dictionaries. In the LHW-E 2000, the headwords within the niche-alphabetical grouping are provided with a grey background to help users find these within the dictionary article. Other dictionaries, such as the PCGW-E4, use one component of the headword (e.g. *Vorsichts-*) to indicate the beginning of a particular niche. Even though nest-alphabetical grouping may help in the learning function for users, it may be more difficult and time-consuming for users to find the looked-for headwords. Consequently, niche-alphabetical arrangements seem to be preferable.

With regard to the microstructures of bilingual dictionary articles, Steinbügl acknowledges the well-known types, namely, integrated, non-integrated and partially integrated microstructures. The different meanings of a headword are usually introduced typographically with alphabet letters or Arabic numerals preceding each separate meaning. In an integrated microstructure, collocations and other multi-lexical units are placed within the polysemic structure, which means that they are presented as part of the information given under each alphabet letter or Arabic numeral. In the case of non-integrated microstructures, multi-lexical units are expanded, which means that they are dealt with in a separate part at the end of the article. Partially integrated microstructures occur when multi-lexical units are sometimes treated in the polysemic structure, and sometimes in an annex designed especially for phraseologisms.

The lexicographical units which have to be dealt with in bilingual dictionary articles differ from those in monolingual dictionaries. In the bilingual dictionary article, each source-language address has to be assigned a target-language item. What makes it difficult, however, is the question whether the dictionary is being compiled for passive foreign language learners, i.e. with a view to understanding foreign language texts, or for active speakers, i.e. with a view to produce texts in the foreign language. Certain information could be superfluous for native speakers, but essential for foreign speakers. Other problems arise with words having multiple meanings, where the presentation of certain data could lead to repetitions. Too much information could also result in users experiencing confusion and despair.

Another problem that she identifies is the way many bilingual dictionaries present collocations under so-called "examples". According to Steinbügl (2005: 25), collocations or syntagms which are placed as examples do not illustrate the use of the headword in question, but rather give the translation of the collocation. An example is the German *mit den Zähnen knirschen*, which can be found under *knirschen* in bilingual dictionaries. This example, being a collocation translated as *grind one's teeth*, does not illustrate the use of *knirschen* (which means "crunch", used in connection with sand, gravel or snow), but in fact represents a figurative meaning of it. Steinbügl (2005: 26) pleads for a micro-structure in active bilingual dictionaries which is rich in collocations, because these multi-lexical units are very important for foreign text production.

She also looks at the arrangement of longer articles with regard to the ease with which users can obtain an overview. Firstly she finds that many bilingual dictionaries base the formation of the articles on semantic criteria, even though the approach is rather arbitrary. When the article for the adjective *ruhig* is observed, for example, it is found to be divided into five semantic groups in the PCGW-E4 and into four in the DOG-E2. Nevertheless, both these articles are fairly clear in their presentation of the information, in contradistinction to the LHW-E 2000 in which the semantic groups are typographically not presented in a user-friendly manner.

Secondly, other dictionaries follow a so-called "categorical-alphabetical principle", such as the Dutch publisher Van Dale's bilingual dictionaries. Here, the noun partner in the collocations is matched with the other partner (which can belong to nine different parts of speech, always numbered from 1 to 9, and always presented in the same order). Steinbügl (2005: 28) thinks that this type of presentation may be too complicated for inexperienced foreign language users, who might not be in a position to classify words according to their parts of speech. She quotes Van der Meer (1998: 222), who states that this type of presentation is a "*lexicographer's* solution to a *lexicographer's* problem and not one tailored to the *user's* problem and the *user's* abilities". Thirdly, in dictionaries without any codes, text blocks or other formal ways of helping users typographically, their searches may be much more difficult and less successful.

In a distinction between monodirectional and bidirectional dictionaries, the latter could have disadvantages because it will always be a compromise to speakers of both language groups. Steinbügl (2005: 35), however, maintains that most bilingual dictionaries are bidirectional. Referring to Hartmann, who states that information in bilingual dictionaries should not be symmetrical, but rather aimed at one particular language group, she claims that the direction of a dictionary can be inferred from the language of the metatexts (e.g. the preface of the DOG-E2 is exclusively in German). She criticises this dictionary by saying that in the articles themselves, the choice of the language for the metatexts is not based on scientific arguments, and when the German verb *stibitzen*, for example, is consulted, the stylistic markers such as "colloquial" given for the target-language equivalents are useful for German, but irrelevant to English

users. She cites many similar examples which result from this type of compromise.

With regard to the functions of bilingual dictionaries, it is important that dictionaries for active use (i.e. the production of texts) should not only give strings of target-language equivalents. Foreign language users are unable to select from these lists the equivalent appropriate for the context. Steinbügl also refers to other scholars such as Mugdan who propagated a series of six bilingual dictionaries for each language set, and who believed that for the production of a foreign language text, users should preferably use a good monolingual dictionary in the foreign language rather than consulting a bilingual dictionary.

After this general discussion of bilingual dictionaries and the various points of view held by scholars, Steinbügl considers collocations as they are found in bilingual dictionaries. She asks the valid question regarding the place where collocations should be listed: under the basis or under the collocator, or under both? The answer will always be influenced by the goal of the dictionary. The best practice is, of course, to list such collocations and their translation equivalents in four places, namely, under the basis as well as the collocator of both language pairs. This is, however, not possible for reasons of space, which means that the lexicographer should try to determine the most likely places where users will look for a particular collocation. Cross-references can then be used, although lexicographers should avoid a too expanded cross-reference system, which may cause confusion and user-unfriendliness.

Sometimes lexicographers use "reduced collocation items". For example, in the case of *Verwirrung stiften*, the PCGW-E4 lists the source-language collocation partner *Verwirrung* under *stiften*, but only presents *cause* as translation equivalent and not the full collocation *cause confusion*. According to Steinbügl (2005: 51), this is very user-unfriendly.

In the third chapter of her book, Steinbügl explains the criteria for her selection of dictionaries to be analysed and the compilation of the corpus to be used in her comparative investigation. Her comparative corpus will consist of collocations of which the translations into English cause problems for German speakers. Discussing the possibilities of statistically investigating collocations, she expresses her concern that the lexicographer's intuition should not be ignored totally. It is possible that machine-driven searches for and analyses of collocations could include word combinations which co-occur accidentally, and by only using such searches, useless results could be obtained. Furthermore, computers cannot (yet) distinguish between the different meanings of linguistic signs. Therefore, she recommends that in the search for collocations, the linguist's intuition should in the first place be trusted, and corpora only used in the second place to confirm this intuition.

Several text corpora are briefly introduced and discussed by means of examples, such as the ones from the Institut für Deutsche Sprache (IDS), the COBUILD Bank of English and the British National Corpus (BNC). A comparison between English and German corpora highlights the different compositions of such databases.

For her comparative corpus, Steinbügl also decided to use monolingual dictionaries for the identification of collocations. Her preliminary corpus consists of 200 collocations, listed on pages 85-89 of her book. Interesting examples of non-equivalence are cited, e.g. *einen Brief einwerfen* or *Staub wischen* (both collocations in German) as opposed to *to post* or *to dust* (no collocations in English). Even the frequency of collocations, when tested against the corpora, sometimes differ rather drastically. For example: The combination *ein Geschenk auspacken* occurs over a hundred times in the IDS, but *to unwrap a present* is found only once in the BNC. When the frequency of *to open a present* is investigated, it appears 31 times, which leads to the conclusion that *to open a present* is a better equivalent for the German collocation. Steinbügl lists many of this type of examples, and discusses some of the problems they cause. Therefore, after taking all kinds of problems into consideration, she sets up a final list of collocations, on pages 103-107, based on scientific argumentation.

The fourth chapter of Steinbügl's book contains her quantitative analysis of the collocational practice in German-English dictionaries. For this purpose, she compares the LGW-DE 1982, the DOG-E2 and the PGW-E1, using graphs to illustrate many results. She looks at the types of information given, the typographical designs, and other features in the presentation of collocations. One of her findings indicates that, in accordance with her expectations, most of the collocations can be found, not under the base, but under the collocator. She discovered that the DOG-E2, which is much more comprehensive than the other dictionaries investigated, gives much more useful information on collocations. This was, of course, also to be expected. The DOG-E2's collocational entries contain many so-called "reduced" collocations, but in its German-English part, the bases are given in English, which is beneficial for German speakers when producing texts. This differs from the LGW-DE 1982, the PGW-E1 and the PCGW-E4, where the German base is given in English under the German collocator, which makes it easier for users to select the right equivalent, but forces them to translate the base into English themselves. When the dictionaries are considered from the point of view of the frequency by which they place collocations under the base, the LGW-DE 1982 ranks the highest and the DOG-E2 the lowest.

In the final chapter, Steinbügl states, however, that her findings are disappointing, because the investigated reference works only treated a relatively small part of the collocations in her comparative corpus comprehensively, or even prominently in bold print. From this, she concludes that there might be a need for special collocational dictionaries.

Steinbügl's book can be very useful for investigating other language pairs from bilingual dictionaries to see what kind of results can be obtained when her approach is used as model. Also, compilers of bilingual dictionaries will do well to note Steinbügl's concern with regard to collocations, and their relatively problematic treatment. Her argumentation, based on linguistic insights and broad experience from her analysis of so many collocations from so many dictionaries, should be taken into cognizance when compiling bilingual dictionar-

ies with other language pairs. The book is written in a German which is not too complicated to be understood by readers with a basic knowledge of German; the examples are clear, and the many graphic illustrations are very helpful.

## References

- Hausmann, F.J.** 1984. Wortschatzlernen ist Kollokationslernen. *Praxis des neusprachlichen Unterrichts* 31: 395-406.
- Hausmann, F.J.** 1985. Kollokationen im deutschen Wörterbuch: Ein Beitrag zur Theorie des lexikographischen Beispiels. Bergenholtz, Henning and Joachim Mugdan (Eds.). 1985. *Lexikographie und Grammatik: Akten des Essener Kolloquiums zur Grammatik im Wörterbuch 28–30.6.1984*: 118-129. Lexicographica. Series Maior 3. Tübingen: Max Niemeyer.
- Van der Meer, G.** 1998. How to Find the Correct Target Language Equivalent. Zettersten, A., V. Hjørnager and J.E. Mogensen (Eds.). 1998. *Symposium on Lexicography VII. Proceedings of the Eighth International Symposium on Lexicography May 2–4, 1996, at the University of Copenhagen*: 25-40. Lexicographica. Series Major 90. Tübingen: Niemeyer.

## Dictionaries

- DOG-E2 = Duden Editors and Oxford University Press (Eds.). 1999. *Duden Oxford. Großwörterbuch Englisch: Deutsch–Englisch/Englisch–Deutsch*. Second Edition. Mannheim: Duden.
- LGW-DE 1982 = Messinger, H. and Langenscheidt Editors (Eds.). 1982. *Langenscheidts Großwörterbuch Englisch. Der kleine Muret-Sanders. Part II: Deutsch–Englisch*. Berlin/Munich: Langenscheidt.
- LHW-E 2000 = *Langenscheidts Handwörterbuch Französisch: Französisch–Deutsch/Deutsch–Französisch*. 2000. Berlin/Munich: Langenscheidt.
- LHW-F 2000 = *Langenscheidts Handwörterbuch Englisch: Englisch–Deutsch/Deutsch–Englisch*. 2000. Berlin/Munich: Langenscheidt.
- PCGW-E4 = Terrell, P., V. Schnorr, W.V.A. Morris and Roland Breitsprecher (Eds.). 1999. *Pons Collins Großwörterbuch für Experten und Universität: Deutsch–Englisch/Englisch–Deutsch*. Fourth Edition. Stuttgart: Klett.
- PGW-E1 = Cop, M. et al. (Eds.). 2002. *Pons Großwörterbuch für Experten und Universität: Englisch–Deutsch/Deutsch–Englisch*. Completely New Edition. Stuttgart: Klett.
- PGW-F1 = *Pons Großwörterbuch: Französisch–Deutsch/Deutsch–Französisch*. 1996. Completely New Edition. Stuttgart: Klett.

Maria Smit  
Department of Music  
University of Stellenbosch  
Stellenbosch  
Republic of South Africa  
(msmit@sun.ac.za)