

BOOK REVIEW

Ying Wang, *The Idiom Principle and L1 Influence: A contrastive learner-corpus study of delexical verb + noun collocations*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2016. xii + 249 pp. ISBN 9789027210746 [Studies in Corpus Linguistics, 77]. <https://doi.org/10.1075/sci.77>

Reviewed by Gaëtanelle Gilquin (Université catholique de Louvain)

Ying Wang's book, entitled "The Idiom Principle and L1 Influence: A contrastive learner-corpus study of delexical verb + noun collocations", is devoted to the use, by Chinese and Swedish learners of English, of verbs that are often assumed to be easy to acquire, namely the high-frequency verbs *have*, *get*, *make*, *do*, *give* and *take*. More particularly, it investigates cases where these verbs are used delexically with a noun collocate (e.g. *have a choice*, *make a statement*, *take a look*). The study relies on the Integrated Contrastive Model (Granger 1996), which seeks to assess the role of mother tongue (L1) influence on the acquisition of a target language (TL) through comparisons between corpus data representing the learners' interlanguage (here, English produced by Chinese and Swedish learners), the target language (here, native English) and the learners' mother tongues (here, native Chinese and Swedish). The starting point of the study is Sinclair's (1991) distinction between the idiom principle and the open-choice principle. The author aims to find out whether learners store and process delexical verb + noun collocations holistically (idiom principle) or whether they build them by filling each slot individually (open-choice principle). However, she considers many other factors as well in her attempt to account for learners' use of the collocations.

The book is made up of seven chapters. The **first chapter** is an introduction which sets the background by introducing some key concepts such as the idiom principle, collocations, high-frequency verbs, learner corpus research and L1 transfer. The three research questions that are addressed in the book are then presented. The first one looks at the frequency of delexical verb + noun collocations in the learner corpora (as compared to the TL and L1 corpora); the second one relies on the same comparisons to investigate the lexical and morphosyntactic features of the collocations; and the third one considers erroneous and unidiomatic uses of the collocations by the learners. The introduction also outlines the Integrated Contrastive Model and briefly describes the teaching and learning context for Chinese and Swedish students of English.

The data and methodology of the study are the focus of the **second chapter**. Six main corpora were involved in the analysis: two learner corpora, the *Corpus of Chinese Learner English* (CCLE) and the *Uppsala Student English Corpus* (USE); two TL corpora, the *Freiburg-LOB Corpus of British English* (FLOB) and the *Freiburg-BROWN Corpus of American English* (FROWN); and two L1 corpora, the *Lancaster Corpus of Mandarin Chinese* (LCMC) and the *Stockholm Umeå Corpus* (SUC). Subsets of each corpus were selected for analysis, consisting of argumentative essays for the learner corpora and of newspaper editorials for the TL and L1 corpora, for a total of about 550,000 words. After describing these (subsets of) corpora, the author explains how the delexical verb + noun collocations in English were identified and extracted, and how the Chinese and Swedish counterparts of the six high-frequency verbs were chosen.

The findings of the study are reported in the next four chapters. Chapter 3 answers the first research question; chapters 4 and 5 answer the second research question; and chapter 6 answers the third research question. The **third chapter** adopts a quantitative perspective by comparing the frequency of the delexical verb + noun collocations (first taken together, and then per verb) in the different corpora, as well as the type-token ratio of the noun collocates. The main result is that the Chinese and Swedish learners tend to overuse the collocations but use less varied noun collocates than the native speakers of English. With respect to the TL and L1 corpora, the Chinese learners appear to come closer to the former, while the Swedish learners combine properties of both. Of the six verbs under study, delexical *have* clearly predominates in all the corpora, quantitatively speaking.

Going beyond this quantitative approach, the **fourth chapter** deals with the noun collocates and how they pattern with the six high-frequency verbs. The chapter starts with a methodological note explaining how collocates were searched for in the *British National Corpus* (BNC), as a complement to the TL corpora, and how all collocates were categorised into different semantic fields. Each of the next sections then focuses on the noun collocates of one verb, presenting the results in the BNC, and then comparing them with the results in the other corpora. The comparison is drawn at the level of the collocates themselves, but also at the level of the semantic categories. Marked semantic preferences are highlighted, some of which are common to learners and native writers, and others are distinctive of one or both learner populations. The examination of the corresponding collocates in the L1 corpora helps assess the possible influence of the mother tongue on learners' usage, which seems to be slightly more marked among the Swedish learners.

While Chapter 4 is devoted to the lexical environment of high-frequency verbs, the **fifth chapter** considers their morphosyntactic environment. It deals with (i) the morphology of the noun collocate, distinguishing between zero-derived nouns (e.g. *change*), derived nouns (e.g. *decision*) and other nouns (e.g. *effort*); (ii) the type of determiner; and (iii) the use of (phrasal or clausal) postmodifiers. In each case, overall results in the different corpora are provided, as well as results per verb. The general trend, visible in both the learner and TL corpora, is that zero-derived noun collocates with a zero determiner and a prepositional-phrase postmodifier are preferred, although there are some differences among the six verbs. Both groups of learners show closer resemblance to the TL than to their respective L1 in terms of these morphosyntactic preferences.

The analyses summarised so far do not distinguish between correct and incorrect uses of the delexical verb + noun collocations produced by the learners. In the **sixth chapter**, by contrast, the focus is on errors and unidiomatic usage. Such uses represent 23% of the collocations produced by the Chinese learners, and 16% of those produced by the Swedish learners. It is the verb *get* that is most often used erroneously or unidiomatically. Both lexical and grammatical problems arise, but the former are more typical of the Swedish learners, whereas the latter are more often found among the Chinese learners. Different types of lexical and grammatical problems are described and their frequencies are compared for each verb in the two learner corpora.

The **final chapter** includes a brief summary of the main findings of the book, centred around the three initial research questions. It also touches upon the pedagogical implications of the study and offers some suggestions for further research, such as a contrastive perspective on the treatment of phraseology in teaching materials or the use of elicitation/experimentation as a

complement to corpus data. The book ends with four **appendices** providing tables with detailed results and an **index** of subjects and authors.

The book offers a comprehensive overview of learners' use of English delexical verb + noun collocations, considering all of their main aspects: their frequency, their lexical and morphosyntactic preferences, and their erroneous and unidiomatic uses. When necessary, the analysis is supplemented with an investigation of the use of noun collocates with other verbs (e.g. *have* vs *get* or *receive education*) or the use of single verbs as alternatives to the collocations (e.g. *take place* vs *happen*). In her attempt to uncover possible traces of L1 influence, the author compares learner data from different L1 backgrounds with target language data, as is common in learner corpus research, but unlike many other studies, these data are also systematically compared with data representing the learners' L1s, thus further substantiating any transfer-related claims. The book does not include any of the fancy statistics that have recently become widespread in the field of learner corpus research. And yet, it does what we would (or should) expect a learner corpus study to do: it provides linguistic insights into learners' actual usage and tries to account for this usage by exploring different interlingual and intralingual factors. What is particularly valuable is that it goes beyond the quantitative descriptions that are typically found in corpus studies to approach the collocations from a truly qualitative perspective. The results are illustrated with many examples taken from the corpora, whose features are examined in detail and in context (including the discursive and socio-cultural context). For instance, the author shows that identical noun collocates may in fact be used differently by the two learner groups (e.g. *benefit* as a collocate of *get* tends to refer to an advantage in the Chinese learner corpus and to social welfare in the Swedish learner corpus); interestingly, she suggests that such differences may be related to non-linguistic factors such as reasoning strategies (e.g. Chinese people's tendency to consider opposite elements, i.e. advantages and disadvantages, before forming a judgment) or economic conditions (e.g. the generous welfare state of Sweden).

The number of factors that are considered to account for the learners' linguistic behaviour is impressive. Next to the central idiom and open-choice principles (which are shown to alternate, sometimes within one and the same collocation) and traditional factors like the L1, the proficiency level, the amount of TL input, stylistic issues and topic-related or teaching effects, the author also looks at aspects that are often neglected in learner corpus research, such as writing conventions and argumentative strategies (e.g. Chinese writers' need for a "concluding moral exhortation" (Matalene 1985: 800)) or cultural differences (e.g. the value of individual freedom in Sweden vs collective well-being in China). In her approach to L1 influence, she takes into account both linguistic and conceptual transfer, as well as perceptual effects (psychotypology), convincingly demonstrating that transfer is neither automatic nor systematic. This, combined with the author's thorough and first-hand knowledge of the Chinese and Swedish languages and cultures, results in a nuanced analysis of the data, which highlights the subtle interplay of factors influencing the learners' use of delexical verb + noun collocations. Admittedly, and as partly recognised by the author herself, it is difficult, or even impossible (at least on the basis of corpus data alone), to ascertain whether these factors are actually at play in the production of a given collocation, and whether this is the case for all the learners who produced this specific collocation (individual variation is not taken into account in the analysis). Moreover, despite the high-frequency of delexical verb + noun collocations, some of the categories that collocations are classified into comprise very low numbers of instances, which do not always allow for reliable generalisations.

The theoretical contribution of the book is not very visible, being diluted throughout the results chapters. Yet, this contribution is very real. After turning the last page, one has not only learned about the use of delexical verb + noun collocations by Chinese and Swedish learners of English, but also about the intricate ways in which L1 influence can materialise as well as the numerous – and sometimes unexpected – factors that may have a linguistic effect. Similarly to Hoey's (1997: 20) quotation, which opens the book, stating that "Even the most unlikely of words is found to have secrets", this book has secrets which are gradually revealed to readers who look beyond the cover and the table of contents. There is more to this book than meets the eye!

References

- Granger, S. 1996. "From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora". In K. Aijmer, B. Altenberg & M. Johansson (Eds.), *Language in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies*. Lund: Lund University Press, 37-51.
- Hoey, M. 1997. "From concordance to text structure: New uses for computer corpora". In B. Lewandowska-Tomaszczyk & P. J. Melia (Eds.), *Proceedings of International Conference on Practical Applications in Language Corpora (PALC'97)*. Łódź: Łódź University Press, 2-23.
- Matalene, C. 1985. "Contrastive rhetoric: An American writing teacher in China". *College English* 47(8), 789-808.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.