

Gilquin, G. 2018. 'Review of Nick C. Ellis, Ute Römer & Matthew Brook O'Donnell, 2016, Usage-Based Approaches to Language Acquisition and Processing: Cognitive and Corpus Investigations of Construction Grammar'. *Applied Linguistics* 39.3: 437-441.

Nick C. Ellis, Ute Römer and Matthew Brook O'Donnell: USAGE-BASED APPROACHES TO LANGUAGE ACQUISITION AND PROCESSING: COGNITIVE AND CORPUS INVESTIGATIONS OF CONSTRUCTION GRAMMAR. Wiley, 2016.

This book, published in the Language Learning Monograph Series, is the result of a long-lasting collaboration between three researchers, Nick Ellis, Ute Römer and Matthew Brook O'Donnell, who come from different – but complementary – backgrounds: cognitive psychology, corpus linguistics and computer science. This combined expertise shines through in each of the ten chapters which, taken together, provide a comprehensive account of the use, acquisition and processing of Verb-Argument Constructions (VACs) in English from a usage-based, and more particularly Construction Grammar perspective. The main focus is on three VACs, namely the Verb Locative, or intransitive motion construction (e.g. *He ran into the room*), the Verb Object Locative, or caused motion construction (e.g. *She put the box under the bed*), and the Verb Object Object, or ditransitive construction (e.g. *They gave me a present*), with occasional references to the Verb Object, or transitive construction (e.g. *I closed the door*). The book has a strong empirical basis and relies on a variety of data, from corpora of child speech or learner language to experimental data coming from free association tasks or lexical decision judgments, among others. These data allow the authors to test a number of clearly identified hypotheses, centred around several general research questions, each of which is the topic of an individual chapter, as briefly outlined below.

Chapter 1 is an excellent introduction to the main themes of the volume, including usage-based theories, Construction Grammar, VACs, the formulaic nature of language or the inseparability of lexis, grammar and semantics. Chapter 2 can also be described as introductory, since it introduces the main factors that are said to affect the learning of constructions. These include frequency, prototypicality, salience and contingency between cue and interpretation. Each of these “determinants of construction learning” (p. 45) is presented with reference to the relevant literature. Chapter 3 corresponds to the first type of analysis of VACs, namely a corpus-based analysis of their use in native English. Using the British National Corpus (BNC), the authors highlight certain characteristics of verbs in VACs: their Zipfian frequency distribution, their selectivity (particular verbs tend to select particular constructions – and vice versa) and their semantic coherence (verbs in a given VAC are organized around networks with prototypical nodes). Because these characteristics are among those that arguably facilitate the learning and processing of constructions, the following chapters examine native and non-native speakers' knowledge, processing and acquisition of VACs and seek to determine “whether speakers are sensitive to the statistics of usage” (p. 99).

The focus of Chapter 4 is on the processing of VACs by adult native (L1) speakers. Free association tasks reveal that, when asked to fill the verb slot in a VAC, subjects are influenced by usage experience as represented in the BNC. More precisely, the verbs that are generated in the experiments tend to correspond to those, in the corpus, with the highest token frequency in the VAC (entrenchment), those that are most faithful to the VAC (contingency) and those that are most central in the semantic network of the VAC (prototypicality). Chapter 5 relies on a similar method, but with non-native (L2) speakers of English. The same three factors turn out to be influential for L2 knowledge and processing, namely entrenchment, contingency and prototypicality. However, the authors also observe effects of L1 interference and of language typology, most notably in relation to Talmy's (2000) distinction between verb-framed and satellite-framed languages, with learners of English with a verb-framed L1 such as Spanish

being less likely to generate verbs that are typical of native English, a satellite-framed language. Chapter 6 is also experimental in nature, but it uses online, hence implicit and automatic, processing tasks, unlike the free association tasks of Chapters 4 and 5 that potentially involve conscious strategies. The five experiments, which include perceptual recognition and lexical decision, largely confirm the results of Chapter 4 for L1 English. This leads the authors to conclude that “words and VACs alike are symbolic representations, acquired from usage, statistics and all, with their subsequent processing tuned probabilistically to usage experience” (p. 184). In Chapter 7, the focus shifts to child language acquisition, with the analysis of data from the CHILDES (Child Language Data Exchange System) database. Characteristics of VACs as they emerged from the previous chapters, such as (near-)Zipfian frequency distribution of the verbs or semantic coherence of the constructions, also appear from the study of child speech and child-directed speech. Children acquire VACs by first learning the frequent and semantically prototypical verbs that are associated with each distinct VAC in the input they receive. A similar mechanism is at work in L2 acquisition, as shown in Chapter 8, with the analysis of corpus data from the ESF (European Science Foundation Second Language) project. The first-learned, pathbreaking verbs in VACs are more frequent and more prototypical in the input, but also more generic and widely applicable. The other slots in VACs mostly follow the same acquisitional route, relying initially on frequent, prototypical and generic items. With Chapter 9, we move to the more computationally oriented section of the volume. Using connectionist and agent-based models, the authors are able to simulate the acquisition of VACs by a language learner from all the input evidence available as well as the transmission of the constructions across generations of speakers. Finally, Chapter 10 brings together the main findings of the book, focusing on their implications, the limitations of the different analyses and the future avenues for research.

The research that is described in this book started as early as 2009, according to some references mentioned in Chapter 8. It has benefited from the collaboration of other researchers, some of whom have co-authored a chapter (David C. Ogden for Chapter 7) or a section (Krishna Bathina for Section 9.5). It is fundamentally interdisciplinary, with a firm foothold in cognitive linguistics, second language acquisition, corpus linguistics, psycholinguistics and natural language processing, but also references to several other fields such as network science or bioinformatics. Since the reader is unlikely to master all of these fields, this could make for a challenging reading experience. However, while the book is dense, it is still accessible on the whole, with illuminating examples or metaphors to explain certain theories (e.g. the organization of books in a library as a metaphor for Rational Analysis on p. 37), figures of different kinds (including heatmaps and tanglegrams) to represent the results visually, and useful summaries at the end of each chapter of analysis to recap the main findings. The amount of data on which the research relies is impressive. The number of subjects involved in most of the experiments is considerable, reaching some 300 people and even more in certain cases. The corpora also provide large quantities of data, with the BNC, for example, containing over 600,000 tokens of the ‘V preposition n’ patterns under study in Chapter 3. All these data are submitted to a variety of sophisticated techniques of analysis and elaborate statistical treatment. The downside of this wealth of data and analyses is that it makes a meticulous approach and detailed account almost impossible. Thus, the ‘V preposition n’ patterns extracted from the BNC could not possibly be disambiguated manually and hence do not all correspond to a VAC. While the authors have carried out pilot studies to improve the recall and (especially) precision of the searches, they end up with a mean precision of 78% and a mean recall of 53%. And while they have removed from their analyses a few patterns for which they could not achieve sufficient precision, they keep some with an even lower precision. The ‘V as n’ pattern, for example, has a precision of 48%, which means that more than half of the sample from which

generalizations are derived actually does not correspond to VACs. The results that are reported are numerous – so much so that over 100 pages of additional tables and figures are to be found (well hidden) on the publisher’s website in the form of “Supporting Information”. Yet, these correspond to only a selection of the results produced by the different analyses. Very often, the authors have to limit themselves to one or two illustrations, which might leave some readers wanting more. Also, authentic examples reproduced from the data are almost completely lacking (some exceptions are found on p. 76 and p. 236), so that one may at times feel they are losing sight of language as such behind the many tables and figures. Another point that should be emphasized, in order to avoid readers’ disappointment, is that, as pointed out on the back cover, this book consists in “a series of research studies”. This explains why, although they all investigate VACs (from several perspectives), these studies also show differences in design that in some cases may restrict the comparability of the results. The specific constructions under investigation may vary from chapter to chapter, or even within one and the same chapter (the ‘V preposition n’ patterns taken from Francis et al. (1996), for example, include from 17 to 20 different patterns, depending on the analysis). Some of the methods and measures differ across chapters (especially Chapter 8 is different in this respect, as noted by the authors on p. 217). The L1 data sometimes represent British English (cf. use of the BNC) and sometimes American English (cf. students from American universities serving as subjects in some of the experiments). As for the L2 data, they represent different mother tongue backgrounds and cover both EFL and ESL situations. The L1 and L2 data also reflect different periods in time, with some corpus data (BNC, ESF) collected in the 1980s and the experimental data collected up to 35 years later. Such differences, of course, are only to be expected given the wide scope of this highly ambitious project, spanning several years of research and involving a number of collaborators. As it is, the book certainly contains the most thorough treatment of the use, acquisition and processing of VACs to date, and one of the best examples of how close collaboration between researchers from different fields can help push the barriers of knowledge and provide new insights that better reflect the multifaceted nature of language as a complex adaptive system. In the final chapter, the authors unveil some of their research plans to continue investigating VACs. There is no doubt that, after having reached the last page of this book, many readers will want to keep following their progress.

## References

- Francis, G., S. Hunston and E. Manning (eds.).** 1996. *Grammar Patterns 1: Verbs*. HarperCollins.
- Talmy, L.** 2000. *Toward a Cognitive Semantics: Typology and Process in Concept Structuring*. MIT Press.

Reviewed by Gaëtanelle Gilquin  
*Université catholique de Louvain, Belgium*  
 E-mail: [gaetanelle.gilquin@uclouvain.be](mailto:gaetanelle.gilquin@uclouvain.be)