

*Review in ICAME Journal, Volume 38, 2014, DOI: 10.2478/icame-2014-0015*

**Hilde Hasselgård, Jarle Ebeling and Signe Oksefjell Ebeling** (eds.). *Corpus perspectives on patterns of lexis*. Amsterdam: John Benjamins. 2013. 299 pp. ISBN 9789027203632 (hardbound); ISBN 9789027271914 (e-Book). Reviewed by **Lot Brems**, University of Liège.

This volume brings together papers that were originally presented at the 32<sup>nd</sup> ICAME conference in Oslo. The contributions are characterized by a fundamentally lexicogrammatical approach, emphasizing that lexis and grammar are not in dichotomous opposition, but should be seen as a continuum. Knowing the meaning of a lexical item subsumes knowing about the lexical as well as grammatical company it prefers to keep, i.e. its collocational and colligational preferences. Obviously, this kind of research goes hand in hand with corpus analysis.

Harking back to neo-Firthian tradition and Hunston and Francis' Pattern Grammar, the volume also links up with such recent and prolific frameworks as (diachronic) construction grammar, constructionalization (Traugott and Trousdale 2013) and grammaticalization. The volume is divided into four sections which each deal with different aspects of lexical patterning based on corpus research. The first section is called 'Sequence and order' and deals with such issues as the relationship between theory ('order') and observable textual data ('sequence'). **Stubbs** starts off with a paper that investigates whether theory-free induction, or data-driven analysis, is possible, since inductive generalizations always involve some kind of pre-corpus analysis theorizing. The identification of collocational patterning (i.e. Sinclair's co-selection) for instance involves semantic abstractions, and is hence a question of order as much as sequence. Stubbs addresses questions such as how to define collocations and semantic prosody, and what counts as one phrasal unit of meaning, especially in view of the fact that lexicogrammatical units show synchronic variation of various kinds and diachronic shifts. Stubbs illustrates this by means of such examples as the *GO-and-VERB* construction which especially in the past tense has developed grammaticalized uses which no longer express motion but subjectified notions of surprise or irritation, as in "*would you believe it, she went and married him*" (p. 16). He concludes that there is no such thing as pure induction, but that it is

not important where an idea comes from, as long as we put it to the test. He also proposes to abandon the corpus-based/driven distinction. Fundamental types of patterning can be positioned on a scale ranging from sequence to order, i.e. collocation, colligation, semantic preference and semantic prosody, which require increasingly more subjective interpretation by an analyst.

The following contributions all present detailed case studies which echo the theoretical concerns voiced in Stubbs. **Dant** studies the sequencing of sex-determined word pairs in binominals, such as *mom and dad*, *men and women*, *ladies and gentlemen*. She takes into account the following interacting and sequential parameters: (1) the metrical constraint (the noun with the fewest syllables comes first), (2) the family relationship constraint (the feminine term precedes in case constraint 1 is inconclusive) and (3) the power constraint (the masculine term precedes the feminine in case there is no family relationship and constraint 1 is inconclusive). She limits the study to noun pairs separated only by *and* in American English and makes use of Mark Davies's Corpus of Contemporary American English (COCA) for the first half of 2011. The data set incorporates various genres. She argues that the three constraints have to be applied sequentially in the way specified above and can account for practically all of the sequences in her data set. Only one noun pair is unaccounted for, i.e. the relatively infrequent *witches and warlocks*. Since the nouns are in the plural, the metrical constraint does not apply. As it does not express a family relationship, only the third constraint should apply, but it does not. Dant brings in the frequency constraint to explain this one odd case: *witches* is the most frequent item of the two (1,356 attestations versus 26 for *warlocks*) and hence comes first.

**Coffey's** contribution looks into sequential preferences of near-synonymous size adjectives such as *enormous great* and *tiny little*. Based on thesaural works, he extracts a set of BIG adjectives and SMALL adjectives which subsume adjectives that also have non-size uses, such as *handsome in a handsome amount of money*. Rather than calling this phenomenon "stacked modification" (Huddleston and Pullum 2002: 562), based on data from the British National Corpus (BNC), he argues that such combinations as a whole act as one unit of meaning. He ends up by looking into sequences of adjectives from other semantic fields.

In the second section, named 'Competing Constructions', the competition and meaning relationships between constructions with similar meanings are examined. **Blanco-Suárez** compares the diachronic developments of two death-related intensifiers *dead* and *deadly* which are the result of grammaticalization and subjectification processes. On the basis of a collocational study it becomes clear that the intensifiers function in different contexts. The distribution seems

determined partly by semantic considerations. *Deadly* displays more concrete meanings, whereas the more highly grammaticalized *dead* has more abstract ones. In addition there are a number of fossilized uses, e.g. *dead drunk*, and ones in which both forms occur in free variation.

**Bachmann** discusses the diachronic competition between *go-V* and *go-and-V* in American English between 1830 and 2009. The genre was restricted to fiction. Using Hilpert's diachronic construction grammar framework he mainly wants to focus on formal changes in both patterns. It has been argued in the literature that *go-and-V* negatively evaluates some event, as in "*What the hell did you go and do that for?*" (p. 93). Overall *Go-V* gained significantly in frequency in the course of the 20<sup>th</sup> century. However, a more detailed look into the specific grammatical contexts both patterns occur in puts this in perspective. *Go-V* has become very frequent in imperative contexts and became the preferred option in this context from the 1920s onwards. Distinctive and diachronic collexeme analysis helps chart changing preferences in the V2 slot. However, Bachmann also indicates the problems of this type of analysis. He concludes by formulating a number of questions that remain unresolved and can be addressed in future research. Among these is the question of meaning differences, such as the expression of negative evaluation which is typically associated with *go-and-V*. However, it turns out that also *go-V* patterns can express this. Other variables, such as register, need to be looked at, and other related constructions should be taken into account, e.g. *go-to-V* and *V-and-V* structures as such. In this way generalizations could be made not only on the micro-constructural level (i.e. the *go-V* versus *go-and-V* construction), but also on the meso-level (i.e. *V-and-V* patterns as such).

**Rohdenburg** describes competition within a family of constructions, i.e. *cannot choose but* + bare infinitive, *cannot but* + bare infinitive, *cannot help –ing* and *cannot help but* + bare infinitive by looking at their present-day variation and their history. In addition there are some blends. Distributional constraints pertain to regional variation as well as differences in text types, the nature of the non-finite verb and certain types of transitivity structures. He argues that the patterns without *help* are the older ones and have been around since the first half of the 16<sup>th</sup> century. *Help –ing* is first recorded in the OED in the early 1700s and the first attestation of *help but* is in 1894. *Help –ing* is more typical of British English and frequent in fiction, whereas the *but*-patterns associate with expository genres. *Help –ing* prefers the company of verbs such as *notice*, *think* and *feel*. Finally, *help –ing* typically occurs in basic active clauses with animate subjects, whereas the other patterns often prefer derived or more complex clauses. In-depth corpus analysis confirms that there is mainly an

opposition between the *help*–*ing* pattern and the three others. It associates with a higher degree of complexity in terms of genre preferences, range of verbs it combines with and syntactic structures it appears in.

**Lorenz's** contribution discusses *gonna* in spoken American English and argues that because of its high frequency this reduced realization of *going to* has become emancipated, i.e. independent of its source form. Using a multivariate analysis he sets out to investigate the factors that determine the choice between *going to* and *gonna*. These include speaker- and speech-related factors. The speaker's age proved to be the strongest determinant of variation with young speakers preferring the contracted form. This synchronic observation can be seen as an indication that a change is in progress. Other significant factors are the type of meaning expressed. The prediction sense most strongly favours *gonna*, while it is disfavoured in deontic uses. He concludes by saying that there still is a split between the spoken and written language, but speculates that *gonna* may eventually be accepted in more formal registers as well. *Going to* would then be restricted to its non-grammaticalized uses.

**Adam Smith's** contribution argues for the existence of complex prepositions and their grammaticalized status. He reviews previous accounts in grammars and checks ways to measure internal cohesion and degrees of decategorialization. He also notes that some complex prepositions, such as *at (the) risk of*, can be used as aspectual markers when followed by gerundial complements (cf. Petré, Davidse and Van Rompaey 2012).

Section three, 'Emerging Patterns' studies new developments in language in terms of the emergence, rise and decline of some items or the development of polysemy in already existing constructions.

**Renouf** tracks the life-cycle of a word and the changing status of neologisms across time, from birth over assimilation to its possible death (and potential revival) in a UK newspaper corpus spanning the period between 1989 and 2011. Using automatic detection software, a number of neologisms are identified (*Eyjafallajökull*, *Arab Spring*, *Graphene*) which illustrate the stages in the life-cycle of neologisms in general. As a result, a more fine-grained definition of neologism is proposed which offers a methodology to assess the status of a neologism and the odds of it being institutionalized.

Based on the self-expanding Norwegian Newspaper Corpus, **Graedler** examines how gender (i.e. common gender vs. neuter) is assigned to a list of 950 recently borrowed English nouns. Grammatical gender is no longer productive in English, but it is in Norwegian, determining agreement with determiners and adjectives as well as the definite singular suffix. Most Anglicisms turn out to have common gender (i.e. encompassing masculine and feminine). It is also

shown that at present (Norwegian) corpora do not always easily lend themselves to this kind of research.

**Diemer** discusses the emergence of new but non-standard prefix verbs with *in-* and *on-*, such as *inbe*, *ontake* and *incame*, using a web-based corpus of blogs. It is argued that computer-mediated communication has facilitated the creation of prefix verbs, a construction which is generally considered to have been on the decline since 1400. In addition to the medium, Diemer argues that a decrease in syntactic complexity, the possibility of wordplay and the influence of non-native language use may all have motivated the re-emergence of prefix verbs.

The two contributions in the final section tackle the question of how one can study meaning in corpus data and is hence aptly called ‘Correlating patterns and meaning’. Using Pattern Grammar, **Vincent** studies the correlation between modal meaning and particular verbal patterns not typically associated with modality, i.e. certain verbs followed by interrogative clauses, e.g. *ask/think/discover/show/determine whether/if*. A quantitative study shows that only six per cent of the sample could not be considered to have modal meaning, including more traditional categories such as dynamic, deontic and epistemic modality, as well as intention and volition, which have a more problematic status in modality research.

**Goossens** evaluates corpus-based and corpus-driven search methods for onomasiological research on expressing quantity approximation in business discourse in Dutch, French and English. It is argued that combining several methods to search annotated corpora yields the best results and generated lists with single- and multi-word approximators such as *around*, *up to* and *a swathe of*.

In sum, this volume presents a comprehensive overview of corpus-based lexicogrammatical research into patterns of language. It covers descriptive as well as important theoretical and methodological topics to do with pattern recognition, synchronic, diachronic and regional competition between constructions, emerging patterns and linking up meaning with specific patterns.

## References

- Hilpert, Martin 2006. Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2 (2): 243–256.
- Hopper, Paul. 1991. On some principles of grammaticization. In E. C. Traugott and B. Heine (eds.). *Approaches to grammaticalization*. Vol. I, 17–36. Amsterdam: John Benjamins.
- Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

- Hunston, Susan and Gill Francis. 2000. *Pattern Grammar: A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Petré, Peter, Kristin Davidse and Tinne Van Rompaey. 2012. On ways of being on the way: Lexical, complex preposition and aspect marker uses. *International Journal of Corpus Linguistics* 17 (2): 229–258.
- Traugott, Elizabeth Closs and Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.