

BOOK REVIEW

Glynn, D. & Robinson, J. A. (Eds.) 2014. *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Amsterdam/Philadelphia: John Benjamins (viii + 545 pp.)

The aim of this book is both “didactic” and “scientific”. On the didactic front, it wants to “broaden the understanding [...] of state-of-the-art corpus linguistic techniques for the study of conceptual structure in Cognitive Semantics” (p. 1); on the scientific front, it is meant to “advance the state-of-the-art of those techniques” (p. 1) by applying them to a number of problems which illustrate the over-arching issue of polysemy and synonymy. The twofold aim is reflected in the structure of the book, whose first “scientific” section (“Polysemy and synonymy”) contains eleven studies focussing on selected phenomena of polysemy and synonymy, whereas the “didactic” Section 2 (“Statistical Techniques”) contains seven chapters offering introductions to a selection of these techniques.

The opening chapters of the sections, both by editor Dylan Glynn, offer general considerations which contextualise the more specific contributions that follow. According to Glynn, “Cognitive Linguistics is undergoing a paradigm shift” (p. 1). This paradigm shift is visible above all in a change from corpus-exemplified to corpus-driven research (p. 9), from introspective to empirical methods (pp. 20–25), and, most importantly, in a growing sophistication of statistical techniques. With one exception, “the entire body of corpus-driven research” between 1985 and 1999 was “restricted to raw counts” (p. 25). The change from corpus-exemplified to corpus-driven research is certainly a welcome move in the direction of methodical rigour, but the problem of representativeness remains. Glynn is certainly right when he says:

No matter how large a corpus, found data will always be biased towards what is common rather than what is possible in a language. Introspection is a vital methodology for proposing hypotheses about what is possible in a language. (p. 27)

But the question is: common where? Perhaps not in the language as a whole, but in some highly specialised part of it, which could be represented by a small but well-targeted corpus. Introspection or intuition will tell us where to look. Biber et al.’s (1998: 246–53) helpful distinction between representativeness and diversity might also be recalled in this context.

Glynn, while impressively rigorous and clear in his methodology, is engagingly humble in defining the aims of Cognitive Linguistics. Unlike Structuralism and Generativism, it does not claim to make “any predictions that can be falsified”

(p. 17; similarly humble on what statistics can do on p. 312). All that Glynn wants is “an operationalisation of the object of study that either offers stability to the [language] system or a means of capturing the dynamic nature of that system” (p. 13). A careful review of previous literature on these topics shows that most of it was “introspective” (p. 25). But, rather than belittle this tradition, Glynn insists that “empirical analysis needs to treat it as foundational”: it has “freed the study of semantic relations from the notions of discrete senses and context independent semantics” (p. 26).

If we have to dispense with discrete senses and context-independent semantics, the “object of study” is best operationalised in “Langacker’s (1987: 59–60) theory of the entrenched form-meaning pair” (p. 13). This theory rests on the assumption that “through repeated contextualised use, the relation between a concept-function and a form becomes stable” (p. 14). In this view polysemy and synonymy are not marginal irritations to the study of meaning but its central concern.

Glynn is also the author of the opening chapter of the “didactic” section, which addresses not only “those less familiar with the field, but also [...] those already working with corpus-driven methods in Cognitive Semantics” (p. 307). The chapter certainly raises our level of problem-awareness. Beginning with a discussion of Collocation Analysis (CA) and Feature Analysis (FA), he argues against the common view that CA is ‘objective’ while FA is more ‘subjective’. Collocations are often regarded as objective facts (there for all to see, as it were), whereas features are the researcher’s own contribution and therefore subjective. Glynn points out that, first, forms are often polysemous “and only certain uses may be relevant for a given study” (p. 311), which makes subjective selection a necessity. On the other hand, the assignment of many features, especially morphosyntactic ones, is entirely objective. Hence, the subjective/objective divide is not as neat as has been claimed.

The chapters making up the “didactic” section are extremely rich in technical detail, to which a short review can hardly do justice. Brief summaries, eked out with a few highlights, is all that can be offered.

Together with Joost van De Weijer, Glynn is also the author of a beautifully patient and lucid step-by-step introduction to the programming language *R* (*R Core Team 2013*), which takes us from the fundamentals of flat dataframes through contingency tables to scripts and packages. The chapter is an attractive teaser to “those less familiar” with *R*. Practically all contributions to the book make use of *R*, which, being open-source and freely available, can also boast of a lively and helpful user community.

Stefan Gries’s chapter on “Frequency tables: Tests, effect sizes, and explorations” is probably one of the most technical parts of the book. His Introduction offers a brief typology of statistical data plus a most apposite methodological reflection concerning Occam’s razor. The application of the razor requires considerable

statistical sophistication (“The Marascuilo procedure”, pp. 378–81), but in the end it simplifies things for the language-learner. Another good example (factors influencing the choice between *try to* and *try and*) is used to illustrate the treatment of multidimensional tables.

Hilpert’s chapter on “Collostructional Analysis” is equally useful and clear. Using the “*it’s* ADJ to V” construction as his paradigm, he discusses three kinds of collostructional analysis in the order of increasing complexity. It ends on an engagingly modest note: “the hard task is not the running of the numbers, but the development of a research question that the chosen technique can answer” (p. 403).

Dagmar Divjak and Nick Fieller explain the principles of cluster analysis with a stunningly trivial feature: similarities and distances between languages are determined on the basis of the first letter of the first ten numerals. After a succinct and useful distinction between kinds of variables (categorical, ordinal, numerical, p. 409) the method is applied in many different ways to eight Germanic languages and to ten Slavonic languages in rather fewer ways. There is also a list of these numerals in nine Romance languages (Table 1) which is not used in the rest of the contribution. From a list of the Germanic numerals (Table 2), a dissimilarity matrix is generated which uses the eight languages as both column labels and row labels (Table 4). At the intersection, for instance, of row “English” with column “Dutch” we find the figure “7”, because Dutch is dissimilar from English in seven first letters out of ten.

Unfortunately and inexplicably, Table 4 is sadly garbled. The eighth column (for Icelandic) is missing, and all dissimilarities involving German have to be increased by 1: German and Swedish, for instance, are dissimilar not in three first letters but four. This error may well affect the cluster formation which is the topic of Sub-chapter 2.3 and is illustrated in Figures 3 to 6 and 8. The wording has to be cautious because the transformation of the dissimilarity matrix is not explained in great detail. But it is clear that the degree of dissimilarity does affect the clustering “Height”, as the legends of the dendrograms say, although it does not map them one-to-one. The step from matrix to dendrogram, which visualises the data, entails a loss of information. This is why there are alternative clustering algorithms.

Divjak and Fieller are appropriately followed by Glynn on Correspondence Analysis. Like Cluster Analysis, it is an exploratory technique, as Glynn insists: “a tool for finding things, not for establishing their significance or discerning their relevance” (p. 444). The method is illustrated with an example taken from an imaginary language which contains six verbs belonging to two semantic classes. Glynn takes his readers through the stages of “flat data-frame”, “cross-tabulation contingency table”, and biplot. At each stage information is lost, but the loss is compensated by a gain in clarity and structure. The biplot, which unites lexemes and grammatical categories in one diagram, shows the two semantic classes as clearly

distinct clusters (Figure 1). Even this small, imaginary example rests on masses of data far exceeding the limitations of print. Readers wishing to appreciate this reduction work should visit the website given on p. 453.

Following on exploratory techniques, Dirk Speelman ends the book with a chapter on Logistic Regression Analysis, a confirmatory technique. Logistic Regression Analysis (as opposed to Linear Regression Analysis) deals with categorical, usually binary outcomes (“response variables”); it relates them to “predictor variables” which may be categorical as well as numerical. Speelman illustrates this with a basketball player, whose success or failure (binary response) is clearly related to his/her distance from the basket (numerical, even continuous predictor). The corpus linguist, even more than the basketball player, is interested in the overall outcome of many events rather than single events. The categorical responses may be said to result in continuous averages. For example, the frequency of a certain variant (say, innovative vs. conservative) is related to language-users’ age, sex, social status as well as to a number of co-textual factors. Clearly, the mathematics behind this is highly complex, and its basics have to be illustrated with linguistic experiments rather than corpora. It is not for nothing that Speelman’s chapter should be the longest in the book – and a worthy conclusion.

Most contributions to Section 1 make explicit reference to these “didactic” chapters. Two chapters on Dutch show this particularly well. Delorge, Plevoets and Coleman extend the discussion of three-argument constructions in an interesting direction: instead of the familiar verbs of giving, they analyse verbs of dispossession. On the basis of a number of Dutch corpora they identify six constructions to which 14 verbs are attracted. All of them begin with *ont-*, the equivalent of “un-” in English and “ent-” in German. Three verbs are found only or almost only in secundative constructions. The relations between the remaining 11 are visualised by Correspondence Analysis, which yields two tight clusters and one outlier, plus a few verbs occupying an intermediate position. Verbs attracted to the double-object cluster (as opposed to the *aan*-dative) “denote more prototypical events of dispossession” (p. 55). 19th-century data (from the literary journal *De Gids*, 1850–99) show less clustering, indicating a “tendency towards *polarization*” in the 20th century (p. 54, italics in the original). Readers of Glynn’s chapter on Correspondence Analysis will find this contribution a methodologically lucid case study, which also enriches insights from earlier, introspective work (p. 48).

Levshina, Geeraerts and Speelman use Multiple Logistic Regression and classification trees to explore the conceptual differences between the Dutch causatives *laten* and *doen* on the basis of an 8-million-word corpus of Netherlandic and Belgian Dutch newspaper texts (Twente NC and Leuven NC). The corpus, balanced for four subject domains, yielded 5,636 occurrences of *laten* and 1,172 of *doen*. The choice between the two constitutes the response variable. Five predictors are

considered: semantic class of (i) Causer, (ii) Causee (both Animate/Inanimate), (iii) Caused Event (Mental/Non-mental), (iv) (In)Transitivity of Effected Predicate, and (v) Syntax of Causee (Marked/Unmarked). The need for a decision tree is nicely illustrated with a side glance at a statistical pitfall: *laten* is so frequent that we would be right in 82.8% of all cases if we predicted *laten* for all contexts (p. 212)! Statistical modelling on the basis of the predictors raises the correct predictions to about 90%. This may seem a small reward for a huge effort, but the true gain is of course in the subtle insights which the modelling affords. To summarise the decision tree, *laten* is strongly preferred when the Causer is Animate; for *doen* to be preferred, two or even three conditions have to be met (p. 214). A more detailed analysis shows *doen* to be most frequent in the contexts of “affective causation” (p. 215ff.). Perhaps surprisingly, this fact does not show in the decision tree. Predictors 2 and 3, which would be most important for such contexts (Animate Causee, Mental Caused Event), do not meet the criteria of the clustering procedure. The authors conclude that “more sophisticated ways of determining the causee’s role could be helpful” (p. 215). Conceivably, the nature of the corpus (newspaper texts) might also be a factor.

Fabiszak, Hebda, Kokorniak and Krawczak explore “The semasiological structure of Polish *myśleć* ‘to think’” with respect to seven prefixes. They begin with a hypothesis resulting from introspection, which is then subjected to the three techniques just discussed. The clustering which is the outcome of their analysis “only partly correlates with the hypothesis” (p. 245). Drawing on their Correspondence Analysis, they suggest that Aspect and Object Semantics are chiefly responsible for the discrepancy.

Jane Klavan’s “Multifactorial corpus analysis of grammatical synonymy” is the only contribution devoted to a non-Indo-European language. Investigating “The Estonian adessive and adposition *peal* ‘on’”, Klavan begins with a succinct presentation of Estonian grammar. Estonian has two ways of saying that a thing is on another thing: synthetically with the Adessive (an inflective case), analytically with the genitive followed by *peal*. With both constructions expressing the same meaning most of the time, one wants to know what determines the choice. Most strikingly, the Adessive is preferred in contexts of greater complexity, which seems to go against the “presumably universal” (p. 259) tendency to prefer the analytic alternative in such contexts: analytic forms are thought to be less demanding in complex situations. Drawing on a large number of semantic and morphosyntactic factors, Klavan argues convincingly that in this case the inflexional alternative is less demanding because it delivers crucial information earlier than *peal*, which, as a *postposition*, comes only after a possibly long, complex, ambiguous locative phrase.

The remaining six chapters deal with English and will be discussed (almost) in order of appearance. Florent Perek uses the English conative construction for

“rethinking constructional polysemy”. Rather than assume one polysemous construction, he follows Croft (2003) in postulating several “verb-class specific constructions” (p. 71). Relying largely on WordNet, he selects three verb-classes for analysis: verbs of cutting, of pulling, and of striking. Rejecting the notion of constructional polysemy, however, he stops short of assigning a distinct constructional meaning to each of his word-classes; his Table 7 knows only “semantic generalization(s)”. The degree to which a verb is “attracted” to the construction varies strongly within the verb-classes. Most strikingly, “the hyperonym of the semantic class is the most repelled collexeme in each case” because these verbs “do not profile any particular semantic trait that would attract them to the construction” (p. 81). Or, in more lay-friendly parlance, the manner of cutting, pulling, etc. is not part of their lexical meaning (on insights to be gained from repelled items, see also Hilpert, p. 397). The paper is a model of scientific rigour, it offers an “overt operationalization” of its analysis, thus making it falsifiable (p. 73).

Co-editor Justyna Robinson’s contribution stands out from the rest of the book in that it is not really a corpus study. Her “Quantifying polysemy in Cognitive Sociolinguistics” is based on “interviews with 72 speakers [...]. Each of the speakers was asked a series of questions aimed at eliciting the most salient usage of [eight] polysemous adjectives” (p. 90). The object of her study is thus not what speakers actually say (or write), but what they say they say. Interestingly, socio-economic background is represented not only by such variables as education and NSEC (National Statistics Socio-Economic Classification score, taking the values of High, Medium, and Lower), but also by “Postcodes”, which may take three different values according to Property prices (Lower, Middle, Higher).

Importantly, the adjectives (*awesome, chilled, cool, fit, gay, wicked, solid, skinny*) are polysemous not for individual speakers but for the population as a whole. Robinson distinguishes 35 senses, including also “reported” ones, i.e. speakers “indicated that they were aware of a certain use of an adjective but they clearly distanced themselves from using this sense” (p. 91). The 35 senses are submitted to cluster analyses and form three or seven clusters, depending on the method used. The three-cluster analysis is discussed in detail and yields age group and social status as strongest predictors. An intriguing picture is offered by Node 4 in Cluster 3 (Figure 8): of 14 speakers aged 31–60 and “living in the highest and lowest postcodes” (p. 110), 13 indicate high usage of this cluster.

Robinson’s cluster analyses lend strong support to her claim that “not only individual words, such as *awesome*, but whole groups of polysemous adjectives currently undergoing semantic change form usage patterns that can be explained by a similar sociolinguistic distribution” (p. 111). In Table 1 she earmarks some of the senses distinguished as “incoming” and some as “potentially disappearing”. A repeat study in, say, 10 to 15 years should be interesting.

Glynn's "The many uses of *run*" is announced as a "repeat analysis" of Gries' (2006) "The many senses of *to run*" (p. 2 and 117). His aim is not to "challenge" Gries, but to "refine" his methods, theory and results (p. 117). Notwithstanding the substitution of *uses* for *senses*, his is of course a semantic study and cannot dispense with the notion of sense(s). It is based on "500 occurrences of [...] *to run* extracted in even proportions from British English and American English and from online personal journals [...] and conversations" (p. 117). In this way Glynn hopes to demonstrate differences between British and American as well as spoken and written English. The senses distinguished in both studies are taken from the same dictionaries (Gries 2006: 91, Glynn pp. 124f.). Both authors aim at giving these dictionary senses a firm empirical basis in the contexts in which the verb *to run* is observed. Glynn's refinements are meant to "integrate the social dimensions of language [here: dialect and register] for descriptive adequacy" (p. 118). He wants to show, *inter alia*, that senses "extend in space", "copy" and "meet" are "distinctly American" (p. 135). I found illustrations only in Gries (2006), as illustrated in Example (1) and Example (2):

- (1) On my way to the elevator, I ran into Pete (Gries 2006: 64)
 (2) Street car tracks run down the center of Pennsylvania (Gries 2006: 67)

For "copy" we may perhaps create:

- (3) ?I'll run a few more copies of the handout [on the copier]

on the model of:

- (4) running the risk of creating disturbances.

Gries (2006: 72) uses Example (4) to illustrate the sense "to risk" of *to run*.

With their different methodologies, there are of course different results as well, but not too many. Gries (2006: 63) begins with 53 different senses identified on the basis of dictionaries and WordNet and presents them in a "radial network" (Figure 1); by hierarchical cluster analysis he reduces them to a dendrogram with 48 different labels (Figure 2). Glynn gives us two hierarchical cluster analyses (Euclidian and Canberra distance) and two correspondence analyses (binary and multiple). The four analyses yield a total of 25 senses, all of which find a counterpart in Gries' 48. In a system without discrete senses, such differences will not matter very much, and the fact that Glynn's senses seem to form a proper sub-set of Gries' is no doubt reassuring – though perhaps more so to the corpus statistician than to the competent but naïve language-user. To the uninitiated, Examples (5) and (6) will hardly suggest that "ignore" and "become used up" are senses of *to run*:

- (5) Catholics run roughshod over Protestant sensibilities. (Gries 2006: 73)
 (6) It's cigs we run out of not petrol. (Glynn, p. 128)

In short, we are being treated not merely to different senses/uses of *to run* in different contexts, but to (more or less) entrenched form-meaning pairs of which *run* is part. A good dictionary will want to list these pairs.

Guillaume Desagulier, setting out to visualise “distances in a set of near synonyms”, studies the collocational behaviour of the four English moderators *rather*, *quite*, *fairly*, and *pretty*. Drawing on the Corpus of Contemporary American English (COCA), he extracts “all adjectives that occur in the first two slots to the right” of them (pp. 155f.). Listing their “top 10 adjectival collocates” (Table 1), he finds that quite a few collocate with more than one of the four. *Good*, by far the most frequent (to be found almost 380,000 times in COCA) co-occurs with all four moderators. *Simple*, with just over 48,000 tokens taking only 19th rank, is shown together with *rather*, *quite* and *fairly*, but not with *pretty*. Given that *pretty* is easily the most frequent of the four moderators (p. 160), this may be surprising: is *simple* only less attracted by *pretty*, or is it repelled by it (perhaps significantly)? True, Desagulier’s topic (“Visualizing distances”) may suggest an interest in attraction, not repulsion (p. 159), and it is pretty (rather? quite? fairly??) unfair to blame a writer for not offering what he has never promised. But after reading Hilpert (p. 397) and Perek (p. 81) on the merits of repulsion, one may be forgiven for some regret.

Similarly regrettable is the neglect of intonation. Admittedly, it has nothing to do with “attraction between lexical items” (p. 145), but it does create differences (cf. Paradis 1997), which could be visualised as distances. It is also of the greatest practical interest. Speaking up once more for the naïve but competent language-user, he or she will be grateful to learn “that moderators are synonyms to some extent” (p. 161), but rather more so for seeing the extent to which they are not.

In the book’s only diachronic contribution, Shank, Plevoets and Cuyckens promise to trace the changing fortunes of *I think that* and *I think* [zero]. With only two possible outcomes (*that* or zero), their problem is an ideal object for Logistic Regression Analysis, where *tertium excluditur* (cf. Speelman, p. 489). Against previous studies, their “corpus-based multivariate analysis” leads to the conclusion that “there is in fact a diachronic *decrease* in zero complementation” (p. 279, italics in the original). This decrease, however, should be seen against another little-noticed fact: “The zero form is clearly the more frequent form from 1560 to 2016” (pp. 285–6). The claim is based on corpus evidence of almost 520 million words of written and almost 250 million words of spoken English. Dividing the number of zero complementizers by the number of *that*-complementizers given in Tables 5 and 6, we may obtain a “zero/*that* ratio” (which the authors do not use, though). For written texts the ratio rises above 10 only in 1990–2009. But for spoken texts it rises to

20 and even 35. In 1994–2012 it drops to 6.04, falling even below the 10.29 of the contemporaneous written corpus. This is tantalizing in the full sense of the word: there is obvious food for our thought, but the nature of the data does not allow us to quench our hunger.

Under these circumstances, speculation is irresistible: the period of the spectacular hump includes the years 1710–1913, for which the authors cite the Old Bailey Corpus as their only spoken corpus (p. 285). It is quite possible that this apparent interaction between mode and period shows only the conventions of a single genre, the court proceedings. Strictly speaking, these proceedings reproduce professional scribal practice just as much as actual speech; they are ‘speech-related’ rather than ‘spoken’. For a plausible reconstruction of earlier English speech one might consider looking at English fiction and English comedies. A fair part of fiction and drama wants to create an illusion of reality, which requires that the language is not too different from what readers and audiences are familiar with. The invention of the phonograph in 1877 has made it possible to record and reproduce speech, and it has taken the new technology a long time to make its impact on the study of the spoken language. Even more important, the mass of archived recorded spoken material increased thanks to media like the radio and the movie. Of course, these reflections do not affect the authors’ comments on the remaining 8 main effects and 8 interactions (Sub-chapters 4.1 and 4.2), but the closing of the gap between speech and writing is probably due to a change in corpus selection rather than “real” language. Which shows once more: corpus study should be supplemented by intuition.

I have broken the order of appearance in order to be able to conclude with a study dealing with two languages rather than only one. Deshors & Gries use the concept of ‘Behavioral Profile’ (BP) for a “multifactorial assessment of learner language”, focussing on the use of *can* and *may* in the English of French learners. BPs may be compared to the datasets in a residents’ registration office. Basically, these will consist of a number of slots with information about residents’ age, gender, profession, address, perhaps income, etc. In statistical use, these “slots” are variables which may assume various “values” or “levels”.

Similarly, a BP holds information about the “morphosyntactic and semantic” features of a linguistic item, typically a lexeme (p. 186). These features are called ‘ID tags’. The morphosyntactic features (person, number, tense, clause type, etc.) are usually straightforward. Semantic features in the present study include not only the familiar “senses” of the modals under study (epistemic, deontic, dynamic); they also distinguish various (classes of) meanings in the subjects and lexical verbs accompanying them. Some of these categories may not spring to mind when thinking of auxiliaries, such as animate vs. inanimate subjects, verbs of accomplishment/achievement/state/process.

The assignment of levels to the variables is laborious and time-consuming, involving as it does a great deal of manual annotation. Its great advantage is that the input of linguistic analysis does not depend on some “striking” feature which may catch the reader’s attention; it can rely on objective data, ideally open to general inspection (perhaps at a website given as reference). The data can be arranged in tabular form, serving as input for a hierarchical cluster analysis to reveal similarities and dissimilarities. A dendrogram (Figure 2) shows that, in terms of semantic variables, the BP of *pouvoir* is more similar to *can* than to *may*: French learners of English are not in the situation of Buridan’s ass between two equally attractive objects. Taking the use of either *may* or *can* by French learners as the dependent variable, the approach offers a multitude of factors (or combinations of such factors) whose relevance we may not have suspected. For instance, French learners tend to prefer *can* over *may* more strongly than native speakers do in subordinate clauses – even more so in negated clauses or with animate subjects (p. 197). It appears that learners want to avoid or at least reduce complexity: the contexts mentioned by way of illustration increase the complexity of the state of affairs to be verbalised, and *can* is felt to be less complex than *may*, because it is closer to *pouvoir*.

One can only agree with the authors’ conclusion: “learners’ ‘non-nativeness’ manifests itself at all linguistic levels simultaneously” (p. 201). The practical consequence is of course not that all deviations from native usage are errors that an examiner would have to mark. But the insights gained should be of great importance stylistically. In ambitious language-teaching they might even lead to exercises which take the factors into account that are responsible for stylistic non-nativeness. Ironically, that might lead to a higher esteem for translation exercises, which some people still regard as didactically backward: French-English translation exercises could select those factors which may “tempt” the learners into preferring the less complex *can* over the more native-like *may* – and teach them to resist that temptation.

Ending on this somewhat utilitarian note is of course not to imply that the other contributions are not useful. Quantitative, corpus-based studies in polysemy and synonymy are almost automatically useful – especially, as I have hinted, for lexicographers. This volume has the extra merit of raising our methodological awareness. Occasionally, the composition of the corpus, especially in terms of genres, might have been given more thought, but the answer to that can only be more corpus research – with a differently composed corpus. The book is mandatory reading for anyone designing a course in corpus linguistics.

References

- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge: CUP. doi:10.1017/CBO9780511804489
- Croft, B. (2003). Lexical rules vs. constructions: A false dichotomy. In H. Cuyckens, T. Berg, R. Dirven & K. Panther (Eds.), *Motivation in Language: Studies in Honour of Günter Radden* (pp. 49–68). Amsterdam/Philadelphia: John Benjamins. doi:10.1075/cilt.243.07cro
- Gries, St. Th. (2006). Corpus-based methods and Cognitive Semantics. The many senses of *to run*. In St. Th. Gries & A. Stefanowitsch (Eds.), *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis* (pp. 57–99). Berlin/New York: Mouton de Gruyter. doi:10.1515/9783110197709.57
- Langacker, R. (1987). *Foundations of Cognitive Grammar*. Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford UP.
- Paradis, C. (1997). *Degree modifiers of adjectives in spoken British English*. Lund: Lund University Press.
- R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at <http://www.R-project.org/> (last accessed March 2017).

Reviewed by Hans-Jürgen Diller (University of Bochum, Germany)