



LUND UNIVERSITY

Polysemy, Syntax, and Variation. A usage-based method for Cognitive Semantics.

Glynn, Dylan

Published in:
New Directions in Cognitive Linguistics

2009

[Link to publication](#)

Citation for published version (APA):

Glynn, D. (2009). Polysemy, Syntax, and Variation. A usage-based method for Cognitive Semantics. In V. Evans, & S. Pourcel (Eds.), *New Directions in Cognitive Linguistics* (pp. 77-106). John Benjamins Publishing Company. http://76.10.210.42/~glynnnet/Glynn_2009_Polysemy_Syntax_Variation.pdf

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Polysemy, syntax, and variation

A usage-based method for Cognitive Semantics

Dylan Glynn

1. Introduction

The study of polysemy has a venerable tradition in Cognitive Linguistics.¹ Since the pioneering of work of Dirven (1981), Radden (1981), Brugman (1983), Lindner (1983), and Vandeloise (1984), the Lexical Network approach to spatial prepositions has proved to be one of the most important contributions of the cognitive paradigm.² However, at both a theoretical and methodological level, this network approach to sense variation has recently come under fire. This study examines a methodological proposal that answers some of the criticisms that a cognitive approach to polysemy faces. The basic premise is to conserve the network model, but to complement this with another method: a corpus-driven quantified and multifactorial method. Such an approach employs a kind of componential analysis that identifies clusters of features across large numbers of speech events. In other words, rather than analyse the possible meanings of a lexeme, a polysemic network should “fall out” from an analysis that identifies clusters of the cognitive-functional features of a lexeme’s usage. These features do not in any way resemble those of the Structuralist componential analyses, since they are not based on a hypothetical semantic system, but describe instances of real language usage and are based upon encyclopaedic semantics of that language use in context.

This usage-based approach is gaining wide currency in Cognitive Linguistics, attested by the wide range of edited volumes dedicated to the subject (Gries and Stefanowitsch 2006; Stefanowitsch and Gries 2006; Zeschel 2008; Glynn and Fischer in press; Newman and Rice in press; Glynn and Robinson forthcoming). We will refer to this approach as the Quantitative Multifactorial method. The discussion begins by briefly covering the current state of the art of polysemy study in cognitive linguistics, where we underline the need for the implementation of this method and how it may cooperate with existing analytical models. In Section 2, the discussion moves to the description of the Quantitative Multifactorial method, and finally the third section examines a case study of the English lexeme *hassle*.

1. Thanks are due to Tine Breban, Sofie van Gijssel, and Koen Pleveots. All shortcomings are my own.

2. Some variations upon and applications of this approach include Schulze (1988, 1994), Hottenroth (1991), Geeraerts (1992), Cuyckens (1993, 1994, 1995), Boers (1996), Bellavia (1996), and Meex (2001).

2. Semasiology: Vagueness, polysemy, and the lexical network

Two groundbreaking studies demonstrate inherent weaknesses in what had become the standard radial model or lexical network approach to the identification sense variation in Cognitive Linguistics. These studies, Sandra and Rice (1995) and Tyler and Evans (2001), have led to a fundamental reconsideration of the conceptual reality represented by network modelling, bringing the validity of such an approach into question. The first study, by Sandra and Rice (1995), led to a debate that questions the psychological reality of sense distinctions proposed by the method and, through the use of psycholinguistic experimentation, raises serious doubts about the validity of the results.³ The second, more theoretical study by Tyler and Evans (2001) uses the very logic that led to the cognitive modelling of sense networks to demonstrate that the network model is flawed.

Thus far, solutions designed to resolve the shortcomings of the lexical network model decline into three approaches.⁴ Firstly, following the psycholinguistic tests of Sandra and Rice (1995), one may turn to an experimental solution to the unverifiable nature of intuition in sense identification and distinction. To these ends, attempts at developing psycholinguistic methods of testing have been pursued. This approach, examples of which include Cuyckens et al. (1997), Sandra and Cuyckens (1999), and Rice et al. (1999), makes the assumption that psycholinguistic experiments are a more reliable method than intuition for the identification of polysemic structures.

The second and third proposed solutions are analytical rather than methodological. One of these analytical solutions is to posit different “types” of polysemy. For instance, Kleiber (1983, 1999), Herskovits (1988), Vandeloise (1990), Deane (1988, 2006), and Glynn (2003, 2006b, 2006c) stress the need to distinguish functional and/or grammatical polysemy from conceptual polysemy.⁵ The assumption here is that, for example, a meaning extension may be the result of the interaction between the conceptual meaning associated with a lexeme and the meaning associated with a given grammatical category. Or, similarly, a socially determined function, such as implicature, may be the motivation behind a semantic extension.

The other analytical solution, developed by Tyler and Evans (2001, 2003, 2004), Evans and Tyler (2004a, 2004b), and Evans (2004, 2005), is named the Principled Polysemy Model and uses predetermined criteria to constrain the proposal of novel senses. This method improves on previous models precisely because sense distinction criteria are overtly specified. By doing this, the approach takes the bull by the horns and begins with the fundamental

3. The debate, exemplified by Croft (1998), Sandra (1998), and Tuggy (1999) follows from an earlier discussion over vagueness and polysemy, cf. Geeraerts (1993), and Tuggy (1993).

4. Other than the three approaches discussed here, there exist a range of other descriptive models more or less built on or derived from the lexical network model. Taylor (2003b) offers a summary of many of these. The crux of the issue is categorisation. Cf. Geeraerts (1989, 1990, 1997, 2000, 2006b), Lehrer (1990), Taylor (1993, 2003a), Cruse (1995), Lewandowska (2002, 2007), and Zlatev (2003).

5. This idea flows from the current trends in Structuralist linguistics that maintain a distinction between *langue* and *parole*. Cf. Fuchs (1987, 1991), Picoche (1994), Victorri and Fuchs (1996), and Rémi-Giraud and Panier (2003) for examples of contemporary Structuralist approaches to polysemy.

question: upon what grounds do we distinguish senses? However, both these analytical solutions (the principled and functional-conceptual) to solving the quandary of sense distinction suffer from the problems that the original Structuralist approach faced: intuition-based methods of investigation necessarily use intensional hermeneutic means for applying their criteria. For example, no matter how clearly set out, the three distinct sense criteria of the Principled Polysemy approach must be applied to language examples and human judgments must be made as to whether these criteria are met.

Nevertheless, every methodology has its strengths and weaknesses. The recent work using the Principled Polysemy approach has met with success because it focuses on precisely the problem at hand rendering the method's inherent weakness open to scrutiny and thus verification. It is thus that the Principled Polysemy approach seeks to resolve the tricky balance between what semantic variation is inherently associated with a given form and what semantic extensions this form–meaning pair may allow in given contexts. This, of course, brings us to the well-known vagueness versus polysemy debate. Principled Polysemy Modelling of Lexical Networks seeks precisely to elucidate the vague polysemic distinction through the proposal of criteria to distinguish “semantic elaborations” from “sanctioning senses” (Evans 2005: 38–40, 41–45).

A priori, all cognitive–functional research accepts axiomatically that “[word] meaning is highly context-sensitive, and thus mutable” (Evans 2005: 71). The question for a linguist faced with such an unstable object of study is not how to render sense variation more stable, but how to reveal structure in its variation. Zelinsky-Wibbelt (2000) poses this fundamental question for the study of polysemy: “Is polysemy a case of lexical representation or rather ... a case of contextual differentiation?” (Zelinsky-Wibbelt 2000: 144). Her discussion is surely amongst the most level-headed on the subject of how to distinguish entrenched sense variation from context-dependent variation. Theoretically in line with Tyler and Evans's (2001) premise, Zelinsky-Wibbelt phrases the problem as a methodological question: in the description of polysemy, “what should be represented at the level of the lexicon and what should be computed by contextual functions?” (Zelinsky-Wibbelt 2000: 145).

At play here is the role of real-time processing versus learnt-automated structure. Necessarily, the former is how one deals with context-dependent meaning production and the latter with entrenched meaning structure. Just as Tyler and Evans (2001: 726) cite universal cognitive processes as a means for explaining the interaction between entrenched lexical reference and contextual information, Zelinsky-Wibbelt poses her methodological question in the context of the basic cognitive process of construal. Here, she echoes the proposal of Kreitzer (1997) that context-dependent construal is an effective means for explaining much sense variation.⁶ She stresses that speakers “negotiate the reorganization of the same concept in potentially infinitely many ways. Vagueness ... represents the speaker's underlying continuum of knowledge from which their communicative ability of negotiating the relevant boundaries of meaning proceeds” (Zelinsky-Wibbelt 2000: 146). Tyler and Evans (2001: 726) stress this same point, arguing that the standard Lexical Network

6. Obviously, the role of different types of cognition in polysemy is a common theme in the literature. Within Cognitive Linguistics, Deane (1988) was probably the first to stress its importance.

approach to polysemy “fail[s] to distinguish between what is coded by a lexical expression and the information that must be derived from context. [Such approaches] fail to take account of meaning construction as a process.” They also follow Kreitzer and Zelinsky-Wibbelt in the use of construal to explain context-dependent polysemy. This brings us to the obvious conclusion that we should base the description of polysemy, and the structures that organise context-dependent construal of sense, on examples of usage, i.e. on examples of how speakers negotiate this basic cognitive and communicative phenomenon.

A Quantitative Multifactorial method aims to fulfil this role, not by adopting an analytical solution, but through a usage-based approach. In this, instead of identifying different senses, one looks for patterns of usage in terms of relative frequency. Its results may complement psycholinguistic testing to corroborate the hypothetical models of polysemy based on intuition. However, it is important to note here that corpus-driven and psychologically tested results cannot disprove intuitive results. This is because no corpus is large enough to account for every possibility and no set of informants large enough to represent the collective speakers of a speech community. Indeed, as we will see, intuitive methods, such as the Principled Polysemy Model, remain essential. However, usage-based quantitative results can complement the intuition-based study of polysemy by verifying core senses and offering information as to the relative semasiological structure of a lexeme. Since a corpus-driven method is inherently restricted to core senses, it cannot serve as a means for constraining the lexical network model. Nevertheless, depending on the representativity of the corpus, results can be argued to describe the most conceptually salient usages of a lexeme or its prototype structure, as well as capture differences between register and dialect. Moreover, and perhaps most importantly, it may reveal how such factors affect each other in the semantic structure of a lexeme. In other words, for a given register, speech situation, or dialect, one reading of a word may be more salient than another. This last point is one of the main advantages of a quantitative method. The multifactorial nature of language use may be rigorously described where, in intuition-based study, one cannot adequately account for the interaction of the different contextual parameters that affect meaning and usage.

Cognitive Linguistics is a usage-based approach to language (Langacker 1988, 2000) and, as such, must necessarily account for the complexities of language as a social phenomenon. This theoretical tenet means that a cognitive approach must necessarily consider extralinguistic parameters. Geeraerts’ plenary lecture at the 8th International Cognitive Linguistics Conference (published Geeraerts 2005) stressed the inevitability of Social Cognitive Linguistics, a point re-iterated by Croft (this volume) in the opening plenary of the first UK Cognitive Linguistics Conference.⁷ The question is not *if* we need to account for variables such as dialect, sociolect, and register, but *how* our analytical apparatuses can account for this complexity. The advantage of the Quantitative Multifactorial method is that this information is inherent to the analysis.

7. Kemmer’s keynote at the *First International Conference of the Swedish Cognitive Linguistics Association* joins the chorus arguing this point. Some other recent publications to press this argument include Tummers et al. (2005), Geeraerts (2006), Grondelaers et al. (2007), Heylen et al. (2008), and Glynn (in press).

The Quantitative Multifactorial method has, in fact, a well-established tradition in Cognitive Linguistics, but principally in the study of onomasiology and parasynonymy. Indeed, some of the earliest cognitive lexical research was quantitative, multifactorial, and usage-based. The early corpus-driven work of Dirven et al. (1982) and the elicitation-based research of Lehrer (1982) are excellent examples of this method. The approach slowly gained momentum with studies such as those of Zelinsky-Wibbelt (1986, 1993), Rudzka-Ostyn (1989, 1995), Schmid (1993), Geeraerts et al. (1994), Atkins (1994), Lemmens (1998), and Geeraerts (1999).

It is in light of the success of this research that Cognitive Linguistics has recently seen a blossoming in quantitative corpus-driven methodology. The use of such methods in the study of lexical and syntactic parasynonymy now represents an important line of research in Cognitive Linguistics. Amongst others, Fischer (2000), Schmid (2000), Gilquin (2003, 2006), Gries (2003, 2006), Grondelaers and Geeraerts (2003), Glynn (2004b, in press, forthcoming), Newman and Rice (2004, 2006), Heylen (2005), Dmitrieva (2005), Divjak (2006) Divjak and Gries (2006), Lemmens (2006), Wulff (2006, et al. 2007), Janda (2007), Frohning (2008), Gries and Divjak (this volume), Coleman (in press), and Janda and Solovyev (forthcoming), are representative of this movement. Both the syntagmatic and paradigmatic parameters are covered as well as conceptual-functional meaning and social-regional variation in usage.

However, extending this methodology to stand as a programmatic method in Cognitive Semantics faces three fundamental hurdles. These difficulties are analytical rather than theoretical in nature and their solutions lie in methodological development. The next section treats each problem in turn.

3. Frames, syntax, and the social dimension. A model for quantitative analysis

3.1 *Tertia Comparationis* and feature analysis

To date, no study has used such quantitative techniques to describe the polysemy of an abstract concept. This is due to the fact that a quantitative approach to polysemy necessitates a *tertium comparationis* (cf. Lewandowska-Tomaszczyk 1998, 1999; Glynn 2004a, 2006a). For abstract concepts, this is impossible since the designatum is a conceptual construct based in culture, possessing no *Lebenswelt* referent. Without an objective constant as a basis for semantic analysis, it is difficult to develop criteria that are sufficiently rigorous to permit the application of quantitative techniques. This problem of operationalising annotation for semantic features cannot be underestimated.

One solution to this problem lies in Frame Semantics (Fillmore 1985). Following Dirven et al. (1982), Rudzka-Ostyn (1989, 1995), and Fillmore and Atkins (2000), Glynn (2006b, 2006c) proposes a solution that goes some way to solving this problem. The reasoning is that one uses the semantic frame as the constant upon which feature studies may be based. This allows both the vagaries of situation context as well as the complexity of the cultural model and its encyclopaedic semantics to be handled in the analysis. What is more, the different arguments and their relations may be treated as semantic features in

the model. The drawback of this approach is twofold. The semantic frame must be posited *a priori* which leads to the same problems that have always plagued intensional definitions and, secondly, such an analytical model is obviously biased towards verbal forms and the concepts that are profiled by them.

Although there seems to be no way of resolving the first issue, this definitional procedure is “overt”. By positing a semantic frame and identifying the arguments and argument relations, the steps taken to define the concept are specified, thus verifiable and less susceptible to analytical flaws resulting from subjective bias. The second issue is more problematic. Although frame semantic structures are still valid for non-verbal concept profiling, if most of the frame arguments are maximally backgrounded and not linguistically expressed in the utterance, the model's ability to capture semantic structure is limited. Thus, for non-verbal profiles, further *ad hoc* parameters may need to be evoked in semantic analysis. This remains a weak point in the model.

3.2 Syntagmatic and paradigmatic dimensions of polysemy

Although both the syntagmatic and paradigmatic dimensions are covered in cognitive approaches to semasiological structure, the interaction between these structures is still not fully understood. The basic problem is that we have not established how schematic and/or morpho-syntactic semantics and less schematic lexical semantics interact. Within Cognitive Linguistics, one position is that syntactic semantics override or coerce lexical semantics (e.g. Talmy 2000). Another position is that there exists a complex interaction between all the various semantic structures in all degrees of schematicity (e.g. Langacker this volume). Following the research presented in Glynn (2002, 2004b, 2008), we make the assumption that syntactic variation affects a polysemy network, and that its effect cannot be satisfactorily predicted by positing meaning structure associated with grammatical forms and classes *a priori*. We must, therefore, account for this variable as an integral part of semantic description. The ramifications of this final point are important. It means that for a given lemma, or root lexeme, there will be semantic variation depending on its syntagmatic context. In other words, its collocation, grammatical class, and even tense or case will necessarily affect the meaning of the item.⁸

This may seem obvious, but to date, within Cognitive Linguistics, the role of this parameter in meaning description has not been considered. We adopt the solution presented in Glynn (in press): instead of treating the polysemy structure of a lexeme as it is expressed for a single part of speech, each lexeme is treated as a onomasiological field, or set of paronyms. This should allow the investigation to meet up with the current movement in Collostructional Analysis (Stefanowitsch and Gries 2003; Gries and Stefanowitsch 2004).

8. The term ‘lemma’ is used following the parlance of corpus linguistics, where it signifies the range of formal variants of the root lexeme.

3.3 Extra-linguistic variation and meaning as usage

Different people use different words in different situations in different ways. This, it would seem, is an undeniable fact of language. In a usage-based approach to language, we must necessarily account for the extra-linguistic factors that this simple statement entails. In other words, the usage and therefore meaning of a lexeme is different in different situations and this semantic variation, or polysemy, is our object of study. The methodology presented here is an attempt at developing a procedure that accounts for extra-linguistic factors, while it is in keeping with the theoretical tenets of Cognitive Linguistics. The basic assumption is that rather than identify conceptual structure, we identify the *various factors of usage that are a result of the conceptual structure that speakers associate with a given form*. The principal factors are dialectal (regional variation), sociolectal (social variation), and register-specific (medium variation).

It seems that the only way to describe the effect of such factors on usage is through corpus-driven quantitative research. The principle is simple: the co-occurrence of features, relative to given “factor variables”, represents structural tendencies in the use of a form. In other words, clusters of semantic features and/or the absence of features in given linguistic and extra-linguistic contexts, are indicators of the meaning of a word. Biber (1995), one of the most important figures in the application of this method, describes this succinctly. In the following quote, his “communicative functions” could be paraphrased as the conceptualisation associated with a form.

Factor interpretations depend on the assumption that linguistic co-occurrence patterns reflect underlying communicative functions. That is, particular sets of linguistic features co-occur frequently in texts because they serve related sets of communicative functions. (Biber 1995: 115)

In various forms, the work of Dirven et al. (1982), Geeraerts et al. (1994), Fischer (2000), and Schmid (2000) makes this assumption and it is accepted here *a priori*.

This assumption is most important and has strong implications for semantic research generally. Employing this method and accepting this assumption means that instead of positing senses and attempting to distinguish them, we simply identify patterns of usage. Degrees of distinctiveness between these patterns may be treated as a statistical question: relative to a given situation (referent, register, region etc.) what is the probability that a given pattern will be used? Approached in this manner, sense identification and distinction are merely summaries, albeit useful ones, of the multifactorial complexity of real language use.

4. A quantitative multifactorial case study of polysemy: *hassle*

In this section we follow a simple case study that shows how a Quantitative Multifactorial method reveals semantic structures that other methods cannot. We examine the semantic structure of the lexeme *hassle* in British and American English. The Fifth Edition of Shorter Oxford English Dictionary and Webster’s Third International Dictionary define *hassle* as:

verb trans. and intrans. Bother, pester, harass (a person); quarrel or wrangle *over* (something). *noun.* A problem, a difficulty; a quarrel, an argument; fuss, bother. (SOED)
n 1. heated argument : WRANGLE. B: a violent skirmish: FIGHT. 2: a protracted debate: CONTROVERSY. 3 a: a state of confusion or commotion: TURMOIL. B: a strenuous effort: STRUGGLE. *vi* ARGUE, FIGHT, DISPUTE. (Webster's)

Although a semantic analysis should not be compared with a dictionary entry, when people are asked to consult their intuition, these definitions seem adequate. There is no mention of variation between American and British, nor any mention of grammatical variation outside the two verbal forms and two nominal forms (the latter implied in the British definition) of *hassle*. Let us see what a coarse-grained quantitative usage-based and multifactorial investigation reveals.

4.1 The corpus and annotation

The mainstay of corpus-driven research focuses on syntactic structures and to these ends powerful parsing technology exists. However, for lexical semantics, tagged corpora are less essential. The most important features of a corpus for the study of content words are its representivity and sheer size. Content words repeat infrequently and lexical variation is typically sensitive to extra-linguistic factors. These two conditions mean that for a lexical semantic study to capture any degree of semantic subtlety of even the most common usages associated with a given lexeme, the corpus must be large and preferably representative of various types of language and register. It is for these reasons that commercially available corpora are less appropriate for lexical semantic investigations. One alternative is to use the internet as a source from which one may build a corpus. Although the internet suffers from thematic bias in its emphasis on “new technologies” as well as a bias of age and social class, the various media that it includes (such as Internet Relay Chat, Usenet, news press, blog-diaries, etc.), mean that a reasonable range of language types is represented.

The current study uses two corpora. The first was made using commercially available “web spiders” that allow one to download large quantities of internet files of a specific kind and from specified servers. The ability to select servers allows one to be reasonably certain about the origin of the text, which is important for concerns of dialect variation. The second corpus was developed by D. Speelman at the University of Leuven and is made up of data extracted from the LiveJournal on-line diary server. LiveJournal represents possibly the largest blog server currently in existence and it kindly allowed us to extract our text. One important feature of the LiveJournal database is that blog-writers must identify which secondary school they attended along with its address. This allows us to be almost entirely certain as to the dialectal origin of the text.

Despite a reasonable range of language types and topics of discourse, these corpora are not as representative as one would normally wish. In order to account for content bias, both theme (topic of discourse) and register (or language type) are systematically annotated. These two parameters as well as dialect, American English versus British English, make up our extra-linguistic factors. Since, within Cognitive Linguistics, we hold that language is a symbolic pairing of form and meaning, the examples are also

annotated for these two basic linguistic parameters. However, in order to operationalise the annotation and render the feature analysis as objective as possible, these parameters were further broken down into a range of variables. These variables can be summarised as four variable groups, two for the parameter of meaning and two for form. Firstly, for the parameter of form, the morpho-syntactic variables were coded separately from the argument structure. This means that the largely objective task of identifying part-of-speech, tense, and so forth can be kept separate from the more theoretically dependent criterion of argument structure.

For the semantic parameter of the feature analysis, the annotation is again divided. Firstly, the more objective of the two variable groups is that of argument types and their relations. Here information such as animacy versus inanimacy and abstractness versus concreteness, as well as (non) familiarity between actors or their power relations, and so forth, are annotated. This is largely objective, although for the adjectival and nominal profilings, this annotation becomes less insightful. However, for the instantiation of a semantic frame, we must note that although one may still annotate backgrounded participants by looking back in the text, this is not always practically possible. A distinction was maintained in the annotation between overt (i.e. profiled and linguistically expressed in the utterance) and covert (backgrounded and not expressed in the utterance) arguments. When there was doubt as to the nature of an argument or an argument relation, it was not annotated.

The second dimension, or variable group, of semantic annotation was the effect on the patient. It is similar to the stimulus feature in the FrameNet project. This feature is highly subjective and thus its results must be treated with caution. The variable includes twelve reasonably fine-grained distinctions that attempt to capture the “effect” upon the patient, such as a “request” being made of the patient or that the patient “feels imposed upon” or “interrupted”. In total, for both the formal and semantic features, 24 variables were annotated. Some of these, such as the morpho-syntactic tagging, were made up of more than 30 binomial values.

4.2 Techniques for Quantitative Analysis

Once the feature analysis is complete and all data are annotated, we need to search for correlations in the frequency of features, relative to the different variables. The field of statistics has an abundance of analytical techniques open for both exploratory investigation and hypothesis testing. The former is used to look for patterns in the data that may be informative; the second is used to determine if these patterns are significant or merely coincidental for a given dataset. This second step is essential since no dataset, no matter how large, can ever represent the reality of the population, in linguistic terms, the culmination of utterances that make up a language.

The results of the feature annotation take the form of cross-tabulations of frequencies, or contingency tables. This may be rephrased as: how often given features occur relative to the different variables specified. In order to examine these results statistically, we must choose from the wide range of exploratory techniques available for the study of categorical data. Our choice is determined by two factors. Firstly because manual semantic coding is labour intensive, our frequency results are relatively small. Obviously the larger the

dataset, the more reliable the results are, and therefore the more statistical techniques become reliable. Secondly, our data are categorical; they are made up of cross-tabulated frequencies of observed features. In other words, either feature x is present or it is not. However, many of the most powerful statistical techniques are designed for the treatment of continuous data and are not appropriate for our categorical results.

Categorical Principal Component Analysis is one of the possible techniques for the treatment of feature frequency in lexical analysis. However, this is a relatively new technique and has not yet been widely applied to this sort of data. Of the various cluster techniques, Model-Based Cluster Analysis may also be appropriate. This has the advantage of being suited to categorical data but has the disadvantage that the number of clusters must be specified before the analysis. Hierarchical Cluster Analysis is another option and is successfully used by Rice (1999), Divjak (2006, and Gries 2006), Gries (2006), Gries and Divjak (this volume), and Gries and Stefanowitsch (in press). Since we are exploring methodological techniques, it would be useful to examine a different method. The method employed here will be Correspondence Analysis. This technique is amongst the simplest to apply and is suited to categorical data.

All these techniques are designed for exploratory analysis. They are used to look for patterns in the data that *may* be representative of significant structure. However, they do not estimate the probability that a given correlation is statistically significant. Significance is the likelihood that the correlation, or relationship between sets of features and variables, is representative of the language as a whole and not just a coincidence in the dataset. In other words, we need to test to determine the probability that a given pattern observed in the data is a result of real factors and not merely coincidence. Obviously, the smaller the number of examples, the harder it is to be sure that the results are representative of the complexity of language reality.

There are many mathematical tests one may use to determine “statistical significance”. There exist also predictive techniques that examine many different variables simultaneously and even offer information as to the relative importance, or effect, of the different variables on the data. Logistic Regression Analysis and Log-Linear Analysis are probably the most appropriate for semantic research. However, for our current purposes, Correspondence Analysis, combined with certain significance tests, should suffice.

4.3 Analysis. The interplay of formal, semantic, and extralinguistic variables

Firstly, let us begin with the formal variation of the lemma. Although it is no secret that different parts of speech or certain collocations result in semantic variation for a single lemma, this parameter is rarely accounted for in polysemy study. Typically one form is chosen and considered in isolation. This, of course, runs contrary to the tenets of Cognitive Linguistics where the different forms associated with the lexical category are choices available for the different profilings of that category. It follows that we should attempt to account for this variation.

The corpus reveals a range of adjectival, nominal, gerundive, and verbal forms of the lemma *hassle*. The attributive adjectival forms are relatively infrequent. Example (1) is typical:

- (1) Specifically, she is interested in how hassling events influence attitudes toward re-frequent-ing a particular store. <www.uncw.edu/aa/2005-2006/csb.html>

The predicative adjectives are common. Although two possible argument structures are possible, a simple stative and a second that expresses the cause of the ‘hassle’ with an oblique, the former is extremely rare. The vast majority of cases overtly express the Cause, which is introduced by *with*, *by*, *over*, *at*, *because*, *for*, or *due to*.

- (2) I keep an eye on things when I can but I’m well hassled by numerous stuff at the moment. <news:9bp2e4\$1@pegasus.csx.cam.ac.uk>

For practical reasons, we will not examine the adjectival forms in any depth.

The nominal variation is important. Other than the gerund, there are mass nouns and both singular and plural count nouns. Let us look at some frequencies of this form relative to dialect. Firstly, we may conflate the singular and plural forms of the count noun. Using the Binomial Exact Test and the Proportional Chi-Squared Test, it is very improbable that there exists a significant difference relative to each other or relative to the dialect variation. In our comparison, we may also include gerunds. However, the gerundive examples pose certain problems in the annotation of their frame structure. Following the FrameNet project, one may divide the gerund examples into “verbal” and “nominal” examples. Although this may at first seem unnecessary, the examples clearly separate into instances where the gerund is part of an event structure and where it is part of a nominal profiling. Nevertheless, relative to dialect, both the nominal and verbal gerunds behave in the same manner, both being highly associated with American English.

In Table 1, we see that although there are relatively similar frequencies of nominals across the two dialects (179 out of 344 and 198 out of 347 occurrences in the respective dialects), their plexity is far from uniform. The Proportional Test is used to demonstrate that it is highly probable that these differences are significant. The *p*-values are listed in the table. Any figure less than 0.05 should be read as significant. Obviously, the closer to zero, the “more reliable” that degree of significance becomes. So, for example in Table 1, a *p*-value of 2.2e-16 (or 0.000000000000000022) is extremely significant. The difference in the frequency of the gerund is also significant, but less so. The Proportional Test uses the Chi-square algorithm and so becomes unreliable with figures under 10. However, the same test applied to the relative British and American frequencies of the verbal gerundive (UK: 4/161, US: 22/143) gives us a *p*-value of < 0.001. Finally, if the nominal and verbal gerundives are combined to give a purely formal category of “gerund” (UK: 6/344, US: 34/347), the test still gives a *p*-value of < 0.001. Therefore, we can be sure that the gerund is highly associated with American English and highly dissociated with British. The different

Table 1. Nominal frequency relative to dialect

Nominal variation	UK	US	Total	Prop. Test
Mass	144	70	214	2.2e-16
Count	33	116	149	1.30e-12
Nom. gerund	2	12	14	0.02199
Total	179	198	377	

grammatical profiling of mass noun, count noun, and gerund obviously represent different meanings of the lexeme. This is an example of the effect of an extralinguistic factor upon semantic structure. Such extralinguistic concerns cannot be sidelined as question of social variation. They make up part of our understanding of how language is used, and therefore part of the encyclopaedic semantics associated with a given form. Let us turn now to the interplay of formal and semantic factors.

In an effort to capture semantic structure without direct annotation of semantic features that tend to be subjectively determined, we can draw on the model of Frame Semantics and annotate Actor types and Relations. In order to see the difference in usage between the nominal and the verbal examples, we may examine the differences in which Actor types are associated with which grammatical constructions. Several of the most common Actor types include a known or “specified” human (Hum_spec), an unknown or non-specified human (Hum_NtSpec), abstract events (Ab_Evnt), concrete events (Ccrt_Evnt), abstract states-of-affairs (Ab_SoA), and concrete things (Thing). Using Correspondence Analysis, we may investigate the relationship between these different Actor types and the constructions. Four constructions are considered here: simple transitive verbs (Trans), transitive verbs with an oblique argument (Trans-obl), resultatives (Result), and nominals (Nominal). Correspondence Analysis uses a reasonably simple statistical technique to examine relative degrees of association. The resulting plots should be interpreted visually, correlations being depicted by relative proximity. The numbers indicated on the axes are there to help determine this relative proximity.

The first obvious grouping (i) is the association between the nominals and inanimate actors. Here, Cause-Actors that are abstract and concrete events, things, and states-of-affairs are clustered with the nominal profiling. This is contrasted with group (ii) that brings together the three verbal constructions and the animate Cause-Actors types. Although an institution is technically not animate, it is clearly a borderline case. In the data, institutions

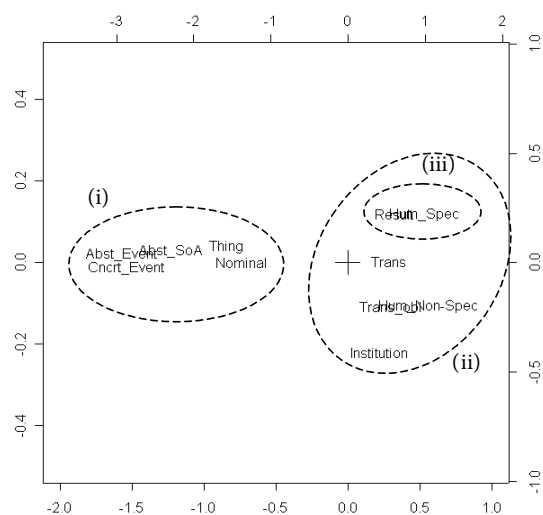


Figure 1. Correlation between Parts of Speech and Actor types

were often whole-for-part metonyms where a bank, university, or some other institution was used to stand for the individuals ‘hassling’ the Patient.

- (3) Debtors have the right to ask collection agencies or any source hassling them for debt collection to stop. <www.public.asu.edu/~hkartadi/laws_in_credit_repair>

More specifically, for the third cluster, the resultative constructions are highly associated with “specified human” Actors. Indeed, they overlap to the extent that the plot is difficult to read. The simple Transitive argument structure is neutral in regard to the specificity of the human feature where the Transitive-Oblique construction is associated with non-specified humans and also institutions. Notice also that known specified human through to institution is depicted as a cline from top to bottom of the group. This suggests that animacy is indeed related to the different Argument structures and grammatical constructions, the resultatives being highly associated with animate Cause-Actors but that this animacy becomes less important for the Cause-Actors of simple Transitive constructions and then less again for the Transitive-Oblique constructions. Let us consider some examples:

- (4) a. since I still have to go through all the hassle of US Immigration
<forum.flightmapping.com/forum_posts.asp?TID=591&get=last>
b. ... my physics exam tomorrow afternoon ... I have been switching rapidly between thinking that it's going to be a piece of piss and thinking I should just top myself now and save the hassle.
- (5) a. Well, I had been hassling Argo for a while to implement my ‘wish list’ into ...
www.heyrick.co.uk/voyager/newsagent/intro.html
b. You were the one hassling me for an answer.
<www.thefridayproject.co.uk/talk/archive/index.php/t-710-p-2.htm>
- (6) a. Some smokers also have a dream that someday the non-smoking world will quit hassling them about their smoking.
<www.nap.edu/books/0309064090/html/169.html>
b. It had me chortling for minutes. If the money's in his wife's account, why aren't they hassling her? <www.thefridayproject.co.uk/talk/archive/index.php/t-2162.html>

Example (4) is typical of the examples captured by the plot in group (i). The choice of the nominal profiling for such events is quite “logical” and we can safely say that non-animate Cause types are suited to the “meaning” expressed by the nominal profiling. Similarly, example (5) represents what seems to be intuitively clear. It is intuitively sound that known human Causes should be common in the resultative examples. What is perhaps less obvious is the animacy cline from familiarity through unfamiliarity to institutional Cause Actors correlating with resultative, simple transitive, and transitive-oblique constructions. This is visible if we compare examples (5) with those in (6).

However, with small frequencies, Correspondence Analysis becomes sensitive to distortions and is less reliable. Although the fact that Transitive-Oblique constructions are highly associated with unfamiliar-human Cause-Actors is intuitively reasonable, it needs further corroboration. It is crucial to remember that this technique is merely an exploratory technique restricted to positing *possible* linguistic structure. Example (6) is offered to show the type of examples that the rather heterogeneous group (iii) represents.

The Intransitive construction was not added to the above Correspondence Analysis because it had low frequencies for all but one of the Actor types. If we look at a couple of the Intransitive examples, we will see why this is a distinct usage.

- (7) a. Officer McCoy, me and him was hassling and my gun went off...
 <www.privacy-council.org.uk/files/other/forrester%20Bowe-rtf.rtf>
 b. It made all the surfers really spread out and we weren't all hassling on the one peak.
 <oneillcwc.asglive.com/daysix>

As we see in the example (7), this usage is semantically distinct from examples (4) to (6). The Intransitive construction is relatively infrequent in the corpus and is semantically marked (despite its prominence in the dictionary entries cited above). This construction draws our attention to an inherent weakness in the quantitative method: infrequent occurrences cannot easily be taken into account. This is a reminder that a corpus-driven quantitative investigation should work in tandem with other methods. The markedness of the Intransitive brings us to the importance of the different constructions associated with the verbal form of the lexeme.

Six basic syntactic forms are revealed. These constructions fall out from the annotation of the different arguments as various semantic roles. In order to capture the different argument structures, the semantic roles of Actor, Cause, Patient, Instrument, and Goal were employed. These were assigned to Subject, Object, and Oblique for all verbal examples. Let us examine the constructions that result from this annotation.

- A. Cause-Oblique Transitive
 Sub. Pred. Obj. Obl.
 She hassled me because of my spots.
 Act. Pred. Pat. Cause.
 ex.: We should all quit hassling the nice oil companies about profits.
 <www.alternate-heaven.com>
- B. Transitive
 Sub. Pred. Obj.
 He hassled me.
 Act./Cause. Pred. Pat.
 ex.: In fact, she hassled him so thoroughly on the street that she made him...
 <www.izzlepaff.com/blog/archives/2004/03/>
- C. Resultative
 Sub. Pred. Obj. Obl.
 She hassled me to eat.
 Act./Ag. Pred. Pat. Goal.
 ex.: If you are caught without the necessary papers/stamp, they hassle you into thinking they will detain... <www.brama.com/travel/messages/4994.html>
- D. Subject-Patient Transitive
 Sub. Pred. Oblique
 He hassled over it.
 Act./Pat. Pred. Cause.
 ex.: Everyone these days is hassling over their weight when they should be hassling about their families. <www.faqfarm.com/Q/How_long_does_it_take_a_12-year-old_to_lose_weight>

- E. Instrumental-Oblique Transitive
Sub. Pred. Obj. *with*-Oblique
He can hassle you with spies.
Act. Pred. Pat. Ag.
ex.: Stop hassling me main man with constant calls! <www.o2.co.uk/services/messaging/voicemail901/celebrityvoicemail>
- F. Intransitive
Sub. Pred.
She hassles
Act. Pred.
ex.: see example (6).

There is, of course, considerably more formal variation than this, especially in the various types of oblique. Differences between *hassle into* something and *hassle to do* something or *hassle over* and *hassle with* represent more than synonymous formal variation. Such variation has clear semantic characteristics. For example, the Cause-Oblique Transitive construction (A) combines with a very wide range of oblique forms, coded with prepositions such as *for*, *because*, *over*, and *on*. It is most likely that such syntactic variation contributes to the polysemy of the lexeme. Here we must, however, restrict ourselves to a coarse-grained investigation.

Firstly, let us see if these forms occur equally across the dialects considered. Again we can use the proportional test to determine whether the differences between the frequencies of each construction are significant.

The results of the proportional test reveal a significant difference between the dialects in the use of the Patient-Oblique Transitive and Resultative Constructions. The latter is especially associated with British English and not American. The other clear difference between the dialects is the use of the Patient-Oblique Construction, which is rare in British but relatively common in American.

Since we have seen that the two dialects are markedly different in their constructional variation for the lexeme, we should investigate how those constructions are associated with different senses for each dialect separately. Once again we can turn to the exploratory technique of Correspondence Analysis for a visualisation of the correlations. If we combine the frequency tables of the constructional variation with one of the semantic variables, we may find correlations between these two variables. The variable of stimulus, or the kind of effect the Cause of the hassle-event has upon the Patient, must be annotated using subjective

Table 2. Dialect variation of construction

Construction	UK	US	Prop. Test
A Cause-Oblique Trans.	10	21	0.02463
B Transitive	110	81	–
C Resultative	34	12	0.003387
D Patient-Oblique Trans.	2	22	1.36e-02
E Instr.-Oblique Trans.	1	4	0.2996
F Intransitive	4	3	–

judgement. However, if results gleaned from its annotation are statistically significant and intuitively plausible, then we may tentatively employ this variable in our analysis.

The plot below represents the results of a Correspondence Analysis of six constructions against six different stimuli. The stimuli considered here are “interrupt-disturb” (interrupt), “repetition-boredom” (rep), “energy-time” (energy), “mental-stress” (mental), “imposition-intrusion” (impose), and “request-solicit” (request). The Intransitive Construction is not included since it was not annotated for stimulus.

Immediately, three correlations are apparent. Firstly, the proximity of the “request” stimulus to the Resultative Construction (C), indicated by (i) on the plot, is a result of what is an intuitively reasonable association. A large percentage of the resultative occurrences describe situations where something is being asked of the patient. Examples (8a)–(8c) are typical of the Resultative Construction–“request” feature correlation.

- (8)
- a.

Be prepared it is pricey and if you’ve booked a table you’ll be given a maximum of 2 hours before the staff start hassling you to have desserts and/or last...
<www.viewlondon.co.uk/info_Pubbar_6564.html>
- b.

You were the one hassling me for an answer, not the other way round
<www.thefridayproject.co.uk/talk/archive/>
- c.

Chad Holleman was hassling Preston for the phone number of some girl he ...
<www.cheddarheads.co.uk/nfle/nfle01/week07.htm>

The relative frequency and importance of this association is clear. Consider below the frequency of occurrences of this construction combined with the “request” feature relative to other semantic features and other constructions. Although by no means a unique correlation, its association is relatively high. Below, we see how its correlation compares with other construction–stimulus feature cross-tabulations.

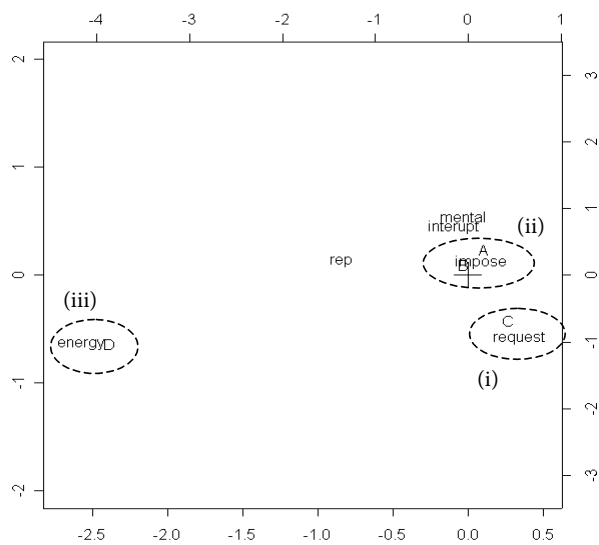


Figure 2. Construction–stimulus correlation in British English

Table 3. Request–resultative correlation in British English

Stimulus feature	Resultative Cx (C) (34 total)	Transitive Cx (B) (110 total)	Cause-Obliq. Cx (A) (10 total)
Request	21	25	1
Impose	18	86	8

The association between the “request” feature and the Resultative Construction relative to the other stimulus features and constructions should be evident.

Secondly, the association depicted in Figure 2 between the stimulus feature of “imposition” with the Transitive construction (B) and the Cause-Oblique Transitive construction (A) is no surprise. These two constructions are the most commonly occurring just as this stimulus feature, “impose”, is the most common. Moreover, this is true for both dialects. It seems plausible that the two basic transitive constructions should be associated with instances of people imposing themselves upon other people. An intuitive reading of the dataset would surely lead a non-quantitative linguistic analysis to propose this sense as a basic sense of verbal *hassle*. Examples (9a)–(9c) are typical.

- (9) a. ... aggressive beggars who profit by hassling members of the public.
 <archive.thisisoxfordshire.co.uk/2001/8/15>
 b. I hate those charities who employ people to hassle you on the streets.
 <www.tiscali.co.uk/forums/showthread.php?t=108045>
 c. However, get there early and there'll be no door staff and no one will hassle you.
 <www.edinburghmetalscene.co.uk/archive/index.php/t-12359.html>

Correspondence Analysis visualises the intuitively sound generality of this sense and identifies its correlation in a quantitative and automated fashion.

The third correlation (iii) is between the stimulus feature of “energy” and the Patient-Oblique Construction (D: *hassle over Pat.*). Due to the low frequency of this construction in British, we must be cautious in reading the plot at this point. However, again this correlation is intuitively valid.

- (10) a. Instead of hassling with multiple drives attached to different computers, you can back it all up to one central location – automatically.
 <www.pixmania.co.uk/uk/uk/183689/art/maxtor/onetouch-iii-shared-stora>
 b. No more hassling with your laptops touchpad or pointing stick.
 <computing.kelkoo.co.uk/>

These examples are typical of those in question and they support the hypothesis that the meaning of the construction fits with the semantics of exhorting energy over something.

Let us turn to the American case. The raw results differ considerably. Let us begin by introducing another three stimulus features that may be relevant. In American English there is a reasonable number of examples where the act of judging the patient negatively is the cause of the state of hassle. These examples were coded as “condemnation”. The two other relevant stimulus features may be referred to as “repetition”, where the patient is hassled due to some repetitive event and “interruption”, where it is an interruption event that has caused the patient hassle.

Table 4. Dialect variation for stimulus feature

Stimulus	British	American	Prop. Test
request	47	23	0.01008
imposition	114	87	0.087
repetition	10	8	–
interruption	18	9	0.1961
energy	6	28	0.000679
condemnation	2	23	7.05e-03
thought	23	30	–

Notice the significant differences between the two dialects on the frequency of the stimulus features of “energy”, “condemnation”, and “request”. “Energy” and “condemnation” are significantly associated with American and not British, contrary to “imposition”, “request”, and “interruption”. Let us plot just the American results here against the constructional variation, once again using Correspondence Analysis.

Again, we see the grouping (i) that is a result of the association between the feature “energy” and Patient-Oblique construction (D). This further verifies the association witnessed for the British data, especially since the data are less sparse for this construction in the American dataset. Similarly, although less common in American, the semantic feature “request” is again highly associated with the resultative construction (ii). Lastly, and also similar to the British data, we have the rough grouping of the more common semantic features and more “basic” grammatical constructions (iii). That is to say, relative to the feature “energy” and the Patient-Oblique construction, as well certain other outliers, these more general semantic features and constructions are clustered. The principal difference between the British and American results here is the outliers. The Intransitive Construction (F) and perhaps also Instrumental-Oblique Transitive construction (E) seem to lack

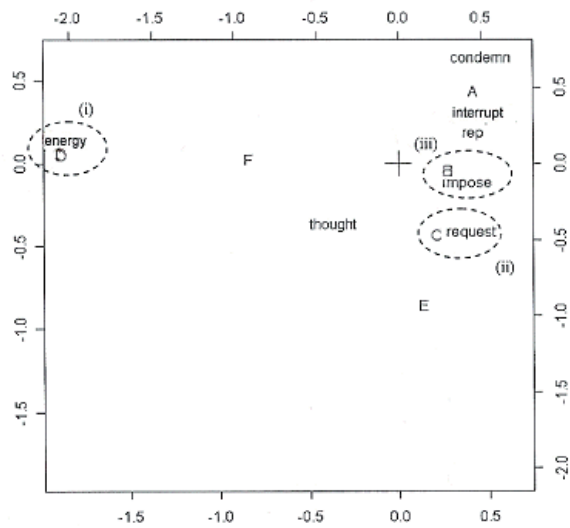


Figure 3. Construction–stimulus correlation in American English

strong associations with any of these semantic features. Likewise, the semantic feature of “thought” is not associated with anything, though it is clearly dissociated from “request” and “condemnation”. This differs from British, where “repetition” was the outlier. Perhaps more data is needed to capture these relations or perhaps they are not particularly associated with any grammatical constructions or any other stimulus features.

Finally, not only does the similarity between the plots in Figure 2 and Figure 3 tell us that the correlations between this semantic variable and the grammatical constructions are relatively similar across the two dialects, having two different datasets serves as a test, adding weight to the hypothesis that these correlations are valid. The next logical step at this point would be to use the dialect as dependent variable and submit the data to a Logistic Regression Analysis. This technique is a confirmatory technique that would allow us to verify this hypothesis. We will not, however, move to confirmatory techniques in this study.

Let us examine another semantic feature. Each example was coded as to whether there was an element of humour involved. Again this feature is subjective in nature and so only very significant differences should be considered reliable. The results here show again a significant difference between the dialects as well as a strong association with one of the constructions.

To begin with, the use of humour and this lexeme is essentially a British characteristic. Out of the 161 British verbal examples, at least 72 were humorous against 17 out of 141 American examples. The proportional test gives a p -value < 0.001 , which should be significant enough to make up for almost any degree of error in the subjective nature of this annotation used. We will, therefore, focus exclusively on the British data for the description of this feature.

We see here that although humour has a high association with both the simple Transitive and Resultative constructions, its association with the Resultative is very significant relative to the overall number of Resultative examples. In other words, although there are more examples of humorous + Transitive co-occurrences, nearly all the Resultatives were humorous.

It must also be remembered that this construction is highly associated with the British dialect as well as the semantic feature of “request”. This was visible in Table 4, above. This shows how the “request” feature is associated with British relative to American, but also that “condemnation” is highly associated with American and not British. Now consider the relations between “humour” and the stimulus features for British.

Table 5. Construction–humour correlation in British English

Construction	+Humour	–Humour	Prop. Test
A Cause-Oblique Trans.	2	8	–
B Transitive	41	69	0.0534
C Resultative	26	8	2.01e-02
D Patient-Oblique Trans.	0	2	–
E Instrumt.-Oblique Trans.	0	1	–
F Intransitive	0	4	–
Total	69	92	

Table 6. Stimulus–humour correlation in British English

Stimulus feature	+Humour	–Humour
request	31	16
imposition	44	70
repetition	5	5
interruption	11	7
energy	1	4

Here we see a strikingly similar correlation to the correlation between the “request” feature and the Resultative construction relative to the “imposition” and the more semantically schematic Transitive construction presented in Table 5. Although there are more humorous examples that possess the stimulus feature of “imposition”, there are almost two times more humorous than non-humorous examples that possess the “request” feature.

This clustering of correlations is beginning to give the picture of semantic structure. The “request” stimulus is associated with the “resultative” construction and both with the humour feature, all three of which are typical of British English and not American. Many would call this a meaning of *hassle*.

We can subject these data to another Correspondence Analysis in order to visualise this correlation. Since this correlation is between three variables, we employ Multiple Correspondence Analysis which follows quite a different procedure but one that is based on similar mathematical principles and whose plots should be interpreted in a similar manner. Figure 4 plots the combinations of the frequencies of constructions and stimulus features against the occurrence and non-occurrence of the “humour” feature.

Firstly, the plot reveals the correlation of the Resultative construction (A.S.C), the stimulus feature of “request” (stim.req), and the humour feature. However, the “humour” feature is, as we know from Table 5, also correlated with the simple Transitive construc-

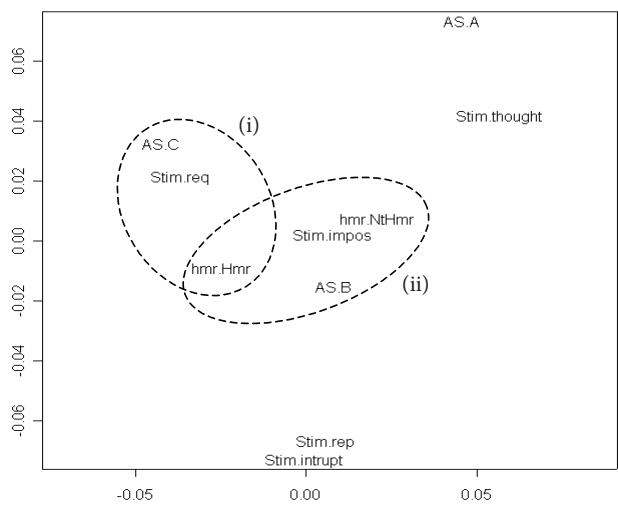


Figure 4. Construction, stimulus, and humour correlation

tion (AS.B) and the “imposition” feature, cf. Table 3. We see this in the proximity of the two groups and the fact that the “humour” feature (hmr.Hmr) is also plotted close to the “imposition” feature and the Transitive. It is for this reason that the groups seem to overlap. The second grouping, of the “imposition” feature (stim.impos), the “-humour” feature (hmr.NtHmr), and the simple Transitive Construction, depicts the kind of usages we saw in example (9), arguably one of the most basic usages of the lexeme.

Importantly, one should note the distance between the features “repetition” (stim.rep) and “interruption” (stim.interrupt) and the constructions Oblique-Cause (AS.A) and Resultative (AS.C). Although we cannot say these stimulus features are highly associated with the simple Transitive construction (AS.B), relative to the other two constructions, there seems to be some association. This is one of the difficulties of reading plots produced in Correspondence Analysis. We cannot say, for instance, that “imposition” is more closely associated with the Oblique-Cause construction than “repetition” or “interruption”, because these latter two are also plotted in association with the Resultative and Transitive constructions. When we look at the frequencies for these correlations this becomes clear. In fact, these two stimulus features had zero co-occurrence with both the Transitive and Resultative constructions, yet 11 (repetition) and 17 (interruption) occurrences with the Oblique-Cause construction. Although small numbers, relative to the 161 British verbal occurrences, this is not negligible.

5. Summary

Although nominals are equally distributed across the two dialects, closer inspection reveals variation. In American, these grammatical classes are generally profiled by gerunds and count-nouns; where in British one finds almost exclusively mass-noun profiling. The semantic variation this almost surely entails was not investigated. However, semantic variation was investigated between the nominal and verbal profilings. A Correspondence Analysis examined the different constructions associated with the lemma and correlated them with Agent types. This revealed what seems to be a clear tendency: the nominal forms are highly associated with Agents that are either events or inanimate things. These Agent types constitute a cline from events through inanimate things to familiar-humans. Verbal forms are at the other end of this continuum and are highly associated with familiar-human Agents. We can describe this difference in usage as polysemic structure. The same analysis discerned another pattern. At the animate end of the continuum, a sub-cline appears, tending from familiar-human through unfamiliar-human to institution. Along this cline of Agent types, the Resultative construction tends to be associated with familiar-humans in contrast to the Transitive and Transitive-Oblique constructions which tend towards “less” animate Agents.

Focusing on verbials, the annotation revealed that six grammatical constructions are associated with the verbal forms of the lemma. Instead of performing a Collostructional Analysis, we continued to investigate the different semantic features associated with each lexeme-construction pairing relative to dialect. Firstly, some construction-pairs seem highly associated with one of the two dialects. For example, the Resultative is essentially

a British usage and the Passive-Oblique Transitive is almost exclusively American. Due to this formal variation between the dialects, the semantic structure was examined for each language variety independently. One of the constructions, the Intransitive construction, was semantically distinct, but its frequency too low to be included a quantitative study.

The semantic analysis focused on two semantic variables. The first of these, the ‘stimulus’ feature, correlates with different constructions in significantly different ways. However, despite the ‘stimulus’ feature variation between constructions and the variation in construction frequency between the dialects, the stimulus–construction associations behave in a similar manner across the dialects. In both cases, the Resultative construction is highly associated with the stimulus of “request” and, importantly, highly disassociated with the other stimulus features and constructions. This adds weight to the first Correspondence Analysis, which suggests that the Resultative is distinct in its association with familiar-human Agents. In addition, a second distinct association was revealed. The Subject-Patient Transitive construction is highly and distinctly associated with the stimulus feature of “energy”. Again this is true of both dialects. Regrettably, this could not be treated in depth. A final correlation brought out by these two Correspondence Analyses is that for both dialects there is a clear association between Cause-Oblique Transitive and the stimulus of “imposition”, where the patient feels put upon by someone or something. Unfortunately, this association also had to be left aside for future investigation.

Finally, we focused on the typically British usage that combines the Resultative construction, the stimulus of “request”, and the familiar-human Agent. To this, we added the second semantic variable, “humour”. Here again, we see a highly significant association with the British usage of the lexeme. Moreover, there is strong correlation between the “humour” feature and the “request” feature, as well as between “humour” and the Resultative construction. The grouping of these features, all of which are distinct relative to other features and forms and specific to British English also coincides with the association of familiar-human Agents that was revealed in the first Correspondence Analysis. Together such a cluster of forms and usages could be reasonably argued to represent a “meaning” of *hassle*.

This possibility was considered in a Multiple Correspondence Analysis. Its results corroborate what is seen in the individual cross-tabulations. However, the analysis suggests that although humour is highly associated with the Resultative–“request” correlation, it is also somewhat characteristic, at least in British, of another more general meaning. In that, the most common construction, the simple Transitive construction, correlates with the most common stimulus feature, “imposition” and this form-meaning pair lies in between the “humour” and “non-humour” features in the results of the Multiple Correspondence Analysis. This finding is intuitively reasonable since the use of humour is generally characteristic of the British dataset.

This raises a final important point. Although these correlations of semantic features such as humour and linguistic forms such as the Resultative can be argued to represent part of the polysemic (semasiological) structure of *hassle*, we must be wary of thematic bias. In other words, these characteristics might be features of the corpus rather than the lexeme in question. Since the kind of language found on the Internet tends to belong to younger speakers and is biased for their topics of discourse, it is possible that this misrep-

resents the importance of this usage in British English. However, the two datasets, British and American, are of the same text type, so we can confidently say that for the quasi-spoken language of the blog-diaries, the differences between the two dialects, relative to this language type, are valid generalisations. Nevertheless, it seems reasonable to assume these findings do represent the language as a whole, but cohort studies with different language types need to be undertaken to verify this.

Last but not least, Correspondence Analysis is only an exploratory method and we are working with relatively small frequencies. The next step is to obtain larger frequencies and attempt to validate the observed patterns using statistical confirmatory techniques such as Log-Linear Analysis and Logistic Regression Analysis. The point of the study was to test the methodology and show that it may reveal semantic structure not detected through intuition-based analyses. In this, the study has proved successful and the integration of results from this kind of investigation with results obtained through a Principled Polysemy analysis should be a straightforward endeavour. Future work needs to operationalise the integration of the Quantitative Multifactorial method and the Principled Polysemy Model for the study of sense variation.

Cognitive Linguistics is a usage-based theory of language and one that assumes language is driven by our encyclopaedic knowledge of the world. In light of this, the kind of usage patterns that Quantitative Multifactorial methods identify offer important clues to the conceptual structures associated with linguistic forms. Although, presenting the results in terms that are typical of the cognitive research community still needs development, mapping the usage, and therefore meaning, of lexemes and constructions is precisely in keeping with the lexical semantic tradition developed by Lakoff (1987). The principal difference is that such quantitative results offer relative tendencies rather than 'different meanings'. This, however, seeing the complex and varied nature of language, is arguably a more cognitively realistic approach to the description of the conceptual structure.

References

- Atkins, B. 1994. Analyzing the verbs of seeing: A frame semantics approach to corpus lexicography. *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society*: 42–56.
- Biber, D. 1995. *Dimensions of Register Variation*. Cambridge: Cambridge University Press.
- Bellavia, E. 1996. The German *über*. In *The Construal of Space in Language and Thought*, M. Pütz & R. Dirven (eds), 73–107. Berlin: Mouton de Gruyter.
- Boers, F. 1996. *Spatial Prepositions and Metaphor: A Cognitive Semantic Journey Along the Up-Down and Front-Back Dimensions*. Tübingen: Narr.
- Brugman, C. 1983. *The Story of Over: Polysemy, Semantics, and the Structure of the Lexicon*. Trier: LAUT.
- Colleman, T. In Press. Beyond the dative alternation: the semantics of the Dutch *aan*-dative. In *Usage-Based Cognitive Semantics. Corpus-Driven Approaches to Conceptual Structure*, D. Glynn & K. Fischer (eds). Berlin: Mouton de Gruyter.
- Croft, W. 1998. Linguistic evidence and mental representations. *Cognitive Linguistics* 9: 151–173.
- Croft, W. This Volume. Toward a social Cognitive Linguistics. In *New Directions in Cognitive Linguistics*, V. Evans & S. Pourcel (eds). Amsterdam: John Benjamins.

- Cruse, A. 1995. Polysemy and related phenomena from a cognitive linguistic viewpoint. In *Computational Lexical Semantics*, P. Saint-Dizier & E. Viegas (eds), 33–49. Cambridge: Cambridge University Press.
- Cuyckens, H. 1993. The Dutch spatial preposition “in”: A cognitive-semantic analysis. In *The Semantics of Prepositions. From Mental Processing to Natural Language Processing*, C. Zelinsky-Wibbelt (ed.), 27–72. Berlin: Mouton de Gruyter.
- Cuyckens, H. 1994. Family resemblance in the Dutch spatial preposition *op*. In *Kognitive Semantik. Ergebnisse, probleme, Perspektiven*, M. Schwartz (ed.), 179–196. Tübingen: Narr.
- Cuyckens, H. 1995. Family resemblance in the Dutch spatial prepositions *Door* and *Langs*. *Cognitive Linguistics* 6: 183–207.
- Cuyckens, H., Sandra, D. & Rice, S. 1997. Towards an empirical lexical semantics. In *Human Contact through Language and Linguistics*, B. Smieja & M. Tasch (eds), 35–54. Frankfurt: Lang.
- Deane, P. 1988. Polysemy and cognition. *Lingua* 75: 325–361.
- Deane, P. 2006. Multimodal spatial representation. On the semantic unity of *over*. In *From Perception to Meaning. Image Schemas in Cognitive Linguistics*, B. Hampe (ed.), 235–284. Berlin: Mouton de Gruyter.
- Dirven, R. 1981. Metaphors of spatial relations. In *Anglistentag Trier 1981*, J. Hasler (ed.), 63–91. Frankfurt: Lang.
- Dirven, R., Goossens, L., Putseys, Y. & Vorlat, E. 1982. *The Scene of Linguistic Action and its Perspectivization by SPEAK, TALK, SAY, and TELL*. Amsterdam: John Benjamins.
- Divjak, D. 2006. Ways of intending: Delineating and structuring near-synonyms. In *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*, St. Th. Gries & A. Stefanowitsch (eds), 19–56. Berlin: Mouton de Gruyter.
- Divjak, D. & Gries, St. Th. 2006. Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2: 23–60.
- Dmitrieva, E. 2005. Об экспансии в русском языке конструкций с предлогом *по*. *Scando-Slavic* 51: 67–83.
- Evans, V. 2004. *The Structure of Time: Language, Meaning and Temporal Cognition*. Amsterdam: John Benjamins.
- Evans, V. 2005. The meaning of *time*: Polysemy, the lexicon and conceptual structure. *Journal of Linguistics* 41: 33–75.
- Evans, V. & Tyler, A. 2004a. Rethinking English ‘prepositions of movement’: The case of *to* and *through*. In *Adpositions of Movement*, H. Cuyckens, W. De Mulder & T. Mortelmans (eds), 247–270. Amsterdam: John Benjamins.
- Evans, V. & Tyler, A. 2004b. Spatial experience, lexical structure, and motivation: The case of *in*. In *Linguistics Studies in Motivation*, G. Radden & K.-U. Panther (eds), 157–192. Berlin: Mouton de Gruyter.
- Fillmore, C. 1985. Frames and the semantics of understanding. *Quaderni di Semantica* 6: 222–254.
- Fillmore, C. & Atkins, B. 2000. Describing polysemy: The case of ‘crawl’. In *Polysemy: Theoretical and Computational Approaches*, Y. Ravin & C. Leacock (eds), 91–110. Oxford: Oxford University Press.
- Fischer, K. 2000. *From Cognitive Semantics to Lexical Pragmatics. The Functional Polysemy of Discourse Particles*. Berlin: Mouton de Gruyter.
- Frohning, D. 2008. From grammatical variation to cognition: Insights from a case study on causal markers in German. In *From Gram to Mind: Grammar as Cognition*, J. R. Lapaire, G. Desagulier, & J. B. Guignard (eds), 167–188. Bordeaux: Presses universitaires de Bordeaux.
- Fuchs, C. 1987. Ambiguïté, vague, polysémie, et continu. *Quaderni di Semantica* 8: 299–305.
- Fuchs, C. 1991. Polysémie, interprétation et typicalité. In *Sémantique et Cognition: Catégories, Prototypes, Typicalité*, D. Dubois (ed.), 161–170. Paris: CNRS.
- Geeraerts, D. 1989. Prospects and problems of prototype theory. In *Prospects and Problems of Prototype Theory*, D. Geeraerts (ed.), 587–612. [Linguistics 27]. Berlin: Mouton de Gruyter.
- Geeraerts, D. 1992. The semantic structure of Dutch *over*. *Leuvense Bijdragen* 81: 205–230.

- Geeraerts, D. 1993. Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics* 4: 223–272.
- Geeraerts, D. 1997. *Diachronic Prototype Semantics. A Contribution to Historical Lexicology*. Oxford: Oxford University Press.
- Geeraerts, D. 1999. Beer and semantics. In *Issues in Cognitive Linguistics*, L. De Stadler & C. Eyrych (eds), 35–55. Berlin: Mouton de Gruyter.
- Geeraerts, D. 2000. Salience phenomena in the lexicon: A typology. In *Meaning and Cognition*, L. Albertazzi (ed.), 79–101. Amsterdam: John Benjamins.
- Geeraerts, D. 2005. Lactal data and empirical variation in Cognitive Linguistics. In *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interactions*, F. Mendoza Ibáñez & M. Peña Cervel (eds), 163–189. Berlin: Mouton de Gruyter.
- Geeraerts, D. 2006. Methodology in Cognitive Linguistics. In *Cognitive Linguistics. Current Applications and Future Perspectives*, G. Kristiansen, M. Achard, R. Dirven & F. J. Ruiz de Mendoza Ibáñez (eds), 3–31. Berlin: Mouton de Gruyter.
- Geeraerts, D., Grondelaers, S. & Bakema, P. 1994. *The Structure of Lexical Variation. Meaning, Naming, and Context*. Berlin: Mouton de Gruyter.
- Gilquin, G. 2003. Causative *get* and *have*. So close, So different. *Journal of English Linguistics* 31: 125–148.
- Gilquin, G. 2006. The place of prototypicality in corpus linguistics: Causation in the hot seat. In *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*, St. Th. Gries & A. Stefanowitsch (eds), 159–192. Berlin: Mouton de Gruyter.
- Glynn, D. 2002. LOVE and ANGER. The grammatical structure of conceptual metaphors. *Cognitive Approaches to Metaphor*. Special edition of *Style* 36: 541–559.
- Glynn, D. 2003. Polysemy and over-polysemy. A three tiered approach to sense variation. Paper presented at the 8th International Cognitive Linguistics Conference. Logroño, Spain.
- Glynn, D. 2004a. Cognitive constructs. Perceptual processing and conceptual categories between mind, language, and culture. In *Language, Culture, and Cognition*, A. Soares da Silva, A. Torres & M. Gonçalves (eds), 247–268. Coimbra: Almedina.
- Glynn, D. 2004b. Constructions at the crossroads. The place of construction grammar between field and frame. *Annual Review of Cognitive Linguistics* 2: 197–233.
- Glynn, D. 2006a. Iconicity and the grammar-lexis interface. In *Insistent Images. Iconicity in Language and Literature*, E. Tabakowska & C. Ljungberg (eds), 267–286. Amsterdam: John Benjamins.
- Glynn, D. 2006b. Concept delimitation and pragmatic implicature. *Perspectives on Metonymy*, K. Kosecki (ed.), 157–174. Frankfurt: Lang.
- Glynn, D. 2006c. Conceptual metonymy – A study in cognitive models, reference-points, and domain boundaries. *Poznań Studies in Contemporary Linguistics* 42: 85–102.
- Glynn, D. 2008. Arbitrary structure, cognitive grammar, and the *partes orationis*. A study in Polish paradigms. In *Naturalness and Iconicity in Linguistics*, K. Willems & L. De Cuypere (eds), 299–333. Amsterdam: John Benjamins.
- Glynn, D. Forthcoming. The semantics of extralinguistic variation. A quantitative study of dialect effects on polysemy. In *Cognitive Sociolinguistics*, G. Kristiansen, D. Geeraerts & Y. Peirsman (eds). Berlin: Mouton de Gruyter.
- Glynn, D. In press. Frames, fields, and synonymy. Developing usage-based methodology for Cognitive Semantics. In *Cognitive Foundations of Linguistic Usage-Patterns*, H.-J. Schmid & S. Handl (eds). Berlin: Mouton de Gruyter.
- Glynn, D. & Fischer, K. (eds). In press. *Usage-Based Cognitive Semantics. Corpus-Driven Approaches to Conceptual Structure*. Berlin: Mouton de Gruyter.
- Glynn, D. & Robinson, J. (eds). In press. *Polysemy and Synonymy. Corpus Methods and Applications in Cognitive Linguistics*. Amsterdam: John Benjamins.
- Gries, St. Th. 2003. *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. London: Continuum.

- Gries, St. Th. 2006. Corpus-based methods and cognitive semantics: The many senses of *to run*. In *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*, St. Th. Gries & A. Stefanowitsch (eds), 57–100. Berlin: Mouton de Gruyter.
- Gries, St. Th. & Divjak, D. This Volume. Behavioral profiles: a corpus-based approach to cognitive semantic analysis. In *New Directions in Cognitive Linguistics*, V. Evans & S. Pourcel (eds). Amsterdam: John Benjamins.
- Gries, St. Th. & Stefanowitsch, A. 2004. Extending collostructional analysis: A corpus-based perspectives on 'alternations'. *International Journal of Corpus Linguistics* 9: 97–129.
- Gries, St. Th. & Stefanowitsch, A. (eds). 2006. *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*. Berlin: Mouton de Gruyter.
- Gries, St. Th. & Stefanowitsch, A. In Press. Cluster analysis and the identification of collexeme classes. In *Empirical and Experimental Methods in Cognitive/Functional Linguistics*, J. Newman & S. Rice (eds). Stanford: CSLI.
- Grondelaers, S. & Geeraerts, D. 2003. Towards a pragmatic model of cognitive onomasiology. In *Cognitive Approaches to Lexical Semantics*, H. Cuyckens, R. Dirven & J. R. Taylor (eds), 67–92. Berlin: Mouton de Gruyter.
- Grondelaers, S., Geeraerts, D. & Speelman, D. 2007. A case for cognitive corpus linguistics. In *Methods in Cognitive Linguistics*, M. Gonzalez-Marquez, I. Mittelberg & M. Spivey (eds), 149–170. Amsterdam: John Benjamins.
- Herskovits, A. 1988. Spatial expressions and the placticity of meaning. In *Topics in Cognitive Linguistics*, B. Rudzka-Ostyn (ed.), 271–297. Amsterdam: John Benjamins.
- Heylen, K. 2005. A quantitative corpus study of German word order variation. In *Linguistic Evidence. Empirical, Theoretical and Computational Perspectives*, S. Kepser & M. Reis (eds), 241–264. Berlin: Mouton de Gruyter.
- Heylen, K., Tummers, J. & Geeraerts, D. 2008. Methodological issues in corpus-based Cognitive Linguistics. In *Cognitive Sociolinguistics. Language Variation, Cultural Models, Social Systems*, G. Kristiansen & R. Dirven (eds), 91–128. Berlin: Mouton de Gruyter.
- Hottenroth, P.-M. 1991. Präpositionen und Objektkonzepte. Ein kognitiv orientiertes, zweistufiges Modell für die Semantik lokaler Präpositionen. In *Approaches to Prepositions*, G. Rauh (ed.), 9–28. Tübingen: Narr.
- Janda, L. 2007. What makes Russian bi-aspectual verbs special. In *Cognitive Paths into the Slavic Domain*, D. Divjak & A. Kochanska (eds), 83–111. Berlin: Mouton de Gruyter.
- Janda, L. & Solovyev, V. Forthcoming. Antonyms as near-synonyms: 'sadness' and 'happiness' in Russian. In *Polysemy and Synonymy. Corpus Methods and Applications in Cognitive Linguistics*, D. Glynn & J. Robinson (eds). Amsterdam: John Benjamins.
- Kemmer, S. 2006. Lexical semantics and corpora: English dimensional terms. Keynote Lecture at the First International Conference of the Swedish Cognitive Linguistics Association. Umeå, Sweden.
- Kleiber, G. 1983. Polysémie et référence: la polysémie, un phénomène pragmatique? *Cahiers de lexicologie* 44: 85–103.
- Kleiber, G. 1999. *Problèmes de sémantique: la polysémie en questions*. Villeneuve-d'Ascq: Septentrion.
- Kreitzer, A. 1997. Multiple levels of schematization: A study in the conceptualization of space. *Cognitive Linguistics* 8: 291–325.
- Lakoff, G. 1987. *Women, Fire, and Dangerous Things. What categories can reveal about the mind*. Chicago: University of Chicago Press.
- Langacker, R. 1988. A usage-based model. In *Topics in Cognitive Linguistics*, B. Rudzka-Ostyn (ed.), 127–161. Amsterdam: John Benjamins.
- Langacker, R. 2000. A dynamic usage-based model. In *Usage Based Models of Language*, S. Kemmer & M. Barlow (eds), 1–63. Stanford: CSLI.
- Langacker, R. This Volume. Constructions and constructional meaning. In *New Directions in Cognitive Linguistics*, V. Evans & S. Pourcel (eds). Amsterdam: John Benjamins.

- Lehrer, A. 1982. *Wine and Conversation*. Bloomington: Indiana University Press.
- Lehrer, A. 1990. Polysemy, conventionality, and the structure of the lexicon. *Cognitive Linguistics* 1: 207–246.
- Lemmens, M. 1998. *Lexical Perspectives on Transitivity and Ergativity. Causative Constructions in English*. Amsterdam: John Benjamins.
- Lemmens, M. 2006. Caused posture. Experiential patterns emerging from corpus research. In *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*, St. Th. Gries & A. Stefanowitsch (eds), 263–298. Berlin: Mouton de Gruyter.
- Lewandowska-Tomaszczyk, B. 1998. Dynamic perspective on antonymous polysemy. In *Making Meaningful Choices in English*, R. Schulze (ed.), 121–137. Tübingen: Narr.
- Lewandowska-Tomaszczyk, B. 1999. The ‘tertium comparationis’ and the concept of equivalence. In *Cognitive Perspectives on Language*, B. Lewandowska-Tomaszczyk (ed.), 53–76. Frankfurt: Lang.
- Lewandowska-Tomaszczyk, B. 2002. Polysemy: Research methodology and mechanisms. In *Cognitive Linguistics Today*, B. Lewandowska-Tomaszczyk & K. Turewicz (eds), 81–96. Frankfurt: Lang.
- Lewandowska-Tomaszczyk, B. 2007. Polysemy, prototypes, and radial categories. In *Handbook of Cognitive Linguistics*, D. Geeraerts & H. Cuyckens (eds), 139–169. Oxford: Oxford University Press.
- Lindner, S. 1983. *A Lexico-Semantic Analysis of English Verb-Particle Constructions*. LAUT: Trier.
- Meex, B. 2001. The spatial and non-spatial sense of the German preposition *über*. In *Polysemy in Cognitive Linguistics*, H. Cuyckens & B. Zawada (eds), 1–36. Amsterdam: John Benjamins.
- Newman, J. & Rice, S. 2004. Patterns of usage for English *SIT*, *STAND*, and *LIE*: A cognitive exploration in corpus linguistics. *Cognitive Linguistics* 15: 351–396.
- Newman, J. & Rice, S. 2006. Transitivity schemas of English *EAT* and *DRINK* in the BNC. *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*, St. Th. Gries & A. Stefanowitsch (eds), 225–260. Berlin: Mouton de Gruyter.
- Newman, J. & Rice, S. (eds). In press. *Empirical and Experimental Methods in Cognitive/Functional Research*. Stanford: CSLI.
- Picoche, J. 1994. A ‘Continuous Definition’ of polysemous items: Its basis, resources and limits. In *Continuity in Linguistic Semantics*, C. Fuchs & B. Victorri (eds), 77–107. Amsterdam: John Benjamins.
- Rice, S. 1999. Aspects of prepositions and prepositional aspect. In *Issues in Cognitive Linguistics*, L. De Stadler & Chr. Eyrych (eds), 225–247. Berlin: Mouton de Gruyter.
- Rice, S., Sandra, D. & Vanrespaille, M. 1999. Prepositional semantics and the fragile link between space and time. In *Cultural, Psychological and Typological Issues in Cognitive Linguistics*, M. K. Hiraga, C. Sinha & S. Wilcox (eds.), 107–127. Amsterdam: John Benjamins.
- Radden, G. 1981. Die übertragenen Bedeutungen der englischen Raumpräpositionen. In *Kassusgrammatik und Fremdsprachendidaktik*, G. Radden & R. Dirven (eds), 133–179. Trier: Wissenschaftlicher.
- Rémi-Giraud, S. & Panier, L. 2003. *La Polysémie ou l'Empire des Sens. Lexique, Discours, Représentations*. Lyon: Press Universitaire de Lyon.
- Rudzka-Ostyn, B. 1989. Prototypes, schemas, and cross-category correspondences: The case of *ask*. In *Prospects and Problems of Prototype Theory*, D. Geeraerts (ed.), 613–661. [Linguistics 27]. Berlin: Mouton de Gruyter.
- Rudzka-Ostyn, B. 1995. Metaphor, schema, invariance: The case of verbs of answering. In *By Word of Mouth: Metaphor, Metonymy, and Linguistic Action in a Cognitive Perspective*, L. Goossens, P. Pauwels, B. Rudzka-Ostyn, A.-M. Simon-Vandenberg & J. Vanparys (eds), 205–244. Amsterdam: John Benjamins.
- Sandra, D. 1998. What linguists can and can't tell you about the human mind. A reply to Croft. *Cognitive Linguistics* 9: 361–378.
- Sandra, D. & Cuyckens, H. 1999. Fuzzy at first sight: The case of two Dutch prepositions. In *Cognition and Function in Language*, B. Fox, D. Jurafsky & L. Michaelis (eds), 200–219. Stanford: CSLI.
- Sandra, D. & Rice, S. 1995. Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language user's? *Cognitive Linguistics* 6: 89–130.

- Schmid, H. J. 1993. *Cottage und Co., idea, start vs. begin: Die Kategorisierung als Grundprinzip einer differenzierten Bedeutungsbeschreibung*. Tübingen: Niemeyer.
- Schmid, H.-J. 2000. *English Abstract Nouns as Conceptual Shells. From Corpus to Cognition*. Berlin: Mouton de Gruyter.
- Schulze, R. 1988. A short story of down. In *Understanding the Lexicon. Meaning, Sense, and World Knowledge in Lexical Semantics*, W. Hüllen & R. Schulze (eds), 395–414. Tübingen: Niemeyer.
- Schulze, R. 1994. Image schemata and the semantics of *off*. In *Kognitive Semantik. Ergebnisse, Probleme, Perspektiven*, M. Schwartz (ed.), 197–213. Tübingen: Narr.
- Stefanowitsch, A. & Gries, St. Th. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8: 209–243.
- Stefanowitsch, A. & Stefan Th. Gries (eds). 2006. *Corpus-based Approaches to Metaphor and Metonymy*. Berlin: Mouton de Gruyter.
- Talmy, L. 2000. *Toward a Cognitive Semantics*. Cambridge MA: MIT Press.
- Taylor, J. 1993. Prepositions: Patterns of polysemisation and strategies of disambiguation. In *The Semantics of Prepositions*, C. Zelinsky-Wibbelt (ed.), 151–175. Berlin: Mouton de Gruyter.
- Taylor, J. 2003a. Cognitive models of polysemy. In *Polysemy – Flexible Patterns of Meaning in Mind and Language*, B. Nerlich, Z. Todd, V. Herman & D. D. Clarke (eds), 31–48. Berlin: Mouton de Gruyter.
- Taylor, J. 2003b. Polysemy's paradoxes. *Language Sciences* 25: 637–655.
- Tuggy, D. 1993. Ambiguity, polysemy, vagueness. *Cognitive Linguistics* 4: 148–162.
- Tuggy, D. 1999. Linguistic evidence for polysemy in the mind. *Cognitive Linguistics* 10: 343–368.
- Tummers, J., Heylen, K. & Geeraerts, D. 2005. Usage-based approaches in Cognitive Linguistics. A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1: 225–261.
- Tyler, A. & Evans, V. 2001. Reconsidering prepositional polysemy networks: The case of *over*. *Language* 77: 724–65.
- Tyler, A. & Evans, V. 2003. *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning, and Cognition*. Cambridge: Cambridge University Press.
- Tyler, A. & Evans, V. 2004. Spatial experience, lexical structure and motivation: The case of *in*. In *Motivation in Grammar*, G. Radden & K.-U. Panther (eds), 157–192. Berlin: Mouton de Gruyter.
- Vandeloise, C. 1984. *L'Espace en Français*. Paris: Seuil.
- Vandeloise, C. 1990. Representation, prototypes, and centrality. In *Meanings and Prototypes. Studies on Linguistic Categorization*, S. Tsohatzidis (ed.), 403–437. London: Routledge.
- Victorri, B. & Fuchs, C. 1996. *La Polysémie: Construction Dynamique du Sens*. Paris: Hermès.
- Wulff, S. 2006. Go-V vs. go-and-V in English: A case of constructional synonymy? In *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*, St. Th. Gries & A. Stefanowitsch (eds), 101–126. Berlin: Mouton de Gruyter.
- Wulff, S., Stefanowitsch, A. & Gries, St. Th. 2007. Brutal Brits and persuasive Americans: variety-specific meaning construction in the *into*-causative. In *Aspects of Meaning Construction*, G. Radden, K.-M. Köpcke, T. Berg & P. Siemund (eds.), 265–281. Amsterdam: John Benjamins.
- Zelinsky-Wibbelt, C. 1986. An empirically based approach towards a system of semantic features. *Proceedings of the 11th International Conference on Computational Linguistics* 11: 7–12.
- Zelinsky-Wibbelt, C. 1993. Interpreting and translating prepositions: A cognitively based formalization. In *The Semantics of Prepositions: From Mental Processing to Natural Language Processing*, C. Zelinsky-Wibbelt (ed.), 351–390. Berlin: Mouton de Gruyter.
- Zelinsky-Wibbelt, C. 2000. *Discourse and the Continuity of Reference. Representing Mental Categorization*. Berlin: Mouton de Gruyter.
- Zeschel, A. (ed.). 2008. *Usage-Based Approaches to Language Processing and Representation*. Special Edition of *Cognitive Linguistics* 19.
- Zlatev, J. 2003. Polysemy or generality? Mu. In *Cognitive Approaches to Lexical Semantics*, H. Cuyckens, R. Dirven & J. R. Taylor (eds), 447–494. Berlin: Mouton de Gruyter.