JEZIKOSLOVLJE
20.1 (2019): 1-30

ČLANCI – ARTICLES – ARTIKEL

UDC 81'367.332.7=111
Original scientific article
Received on 12.12. 2018
Accepted for publication on 15.01. 2019
https://doi.org/10.29162/jez.2019.1

Karolina Krawczak[1]
Dylan Glynn[2]
[1]Adam Mickiewicz University, Poznań
[2]Université Paris 8 Vincennes-Saint-Denis

Operationalising construal. A corpus-based study in cognition and communication constructions

This study seeks to demonstrate that the Behavioural Profile Approach, specifically Multifactorial Usage-Feature Analysis (Geeraerts et al. 1994; Gries 2003), can be used to quantitatively describe lexico-grammatical construal (Langacker 1987; 1999). It examines the *of – about* constructional alternation for the complementation of cognition and communication predicates. The predicates sampled include *know*, *speak*, *talk*, and *think* distributed across the two prepositions in British and American English. In total, a sample of some 700 occurrences are analysed; the annotation schema is based on previous literature in the field (Radden 1981; Rudzka-Ostyn 2003; Dirven 2003; Lindstromberg 2010). Using a combination of mixed-effects logistic regression, multiple correspondence analysis, and loglinear analysis, the study is able to successfully identify a behavioural profile of the two alternations, which can be interpreted as an operationalisation of the opposing construals. Although distinct profiles are obtained, an adequate means for separating the conceptual contribution of the predicate and the complement will require further investigation.

Key words: predicates; prepositional complementation; construal; corpus; multivariate statistics.

1. Introduction

Conceptual construal, or the ability to conceive the same object from different perspectives, is held to be a fundamental cognitive ability and one that is a cornerstone of Cognitive Grammar. Different expressions, which are in an onomasiological relation, are believed to profile different socio-functional or perceptual-conceptual views of a scene. The ability of language to foreground and background different ways of thinking about the same thing is seen as not only fundamental to language structure or grammar, but also to the negotiated communicative process.

The theory of construal finds its origins in Gestalt psychology, which sought to understand how we cognise and categorise our perceived world. Cognitive Linguistics continues this practice, yet adopts a behaviourist usage-based model of grammar (Langacker 1987; 1988). Despite important advances in developing analytical apparatuses for identifying the different types of construals that language encodes, operationalising the notion in frequency-based terms that will inform usage-based descriptions of language structure has yet to be systematically achieved. This study seeks to operationalise the notion in a way that can be applied to the analysis of observational data. A frequency-based study of how construals are employed in language will permit the use of observational evidence for testing the predictive power and descriptive accuracy of the construal types already proposed through experimental and introspective methods.

The constructional alternation of the prepositional complements *of* and *about* is examined relative to their use with two mental predicates *think* and *know* and two communication predicates *speak* and *talk*. The aim is to identify usage patterns of construals associated with two prepositional complements that are not a result of the lexical semantics associated with the predicates.

1.1. *Construal and lexicogrammatical profiling*[1]

Making sense of the surrounding world is a matter of conceptualization. This is a process whereby the subject approaches the "object of conception", to use Langacker's terminology, from a given psycho-somatic perspective, and, vis-à-vis the relevant aspects of his/her experientially incremented knowledge as well as the immediate contextual information, construes it in a certain manner. This construal or 'portrayal' is imposed upon the pertinent conceptual content, activated contextually, thus engendering meaning (Langacker 1999: 205; 2013: 43). Generally speak-

[1] This section is based on the following sources: Langacker (1999; 2008; 2009; 2013).

ing, therefore, construal "refers to our manifest ability to conceive and portray the same situation in alternate ways" (Langacker 2013: 43). This relationship between the conceptualiser and the object being conceptualised is established and realized along lexical and syntactic lines. Langacker (2008: 55–85; 2009: 6–10) identifies a range of such lexico-grammatical parameters that determine construal, among which the following can be listed: (1) the level of specificity, also referred to as granularity or resolution; (2) the perspective from which the conceptualised situation is viewed (e.g., vantage point, subjectivity); (3) prominence concerning such matters as profiling and figure/ground distinction; and (4) focusing.

The first criterion has to do with how specifically or schematically the speaker chooses to present the situation under conceptualization. This can be exemplified by such paradigmatically related lexical items as: *thing* > *piece of furniture* > *sofa* > *comfortable sofa*. Depending on the communicative context of use, the speaker may select one of the above levels of categorization to refer to one and the same object. The second construal parameter, concerning the perspective adopted to conceptualise a given scene, affects such aspects of meaning as vantage point, subjective vs. objective construal or dynamicity of the portrayal. Vantage point, by default identified with the position of the interlocutors, is inherent in many, if not all, expressions, as illustrated by lexemes such as *upstairs* vs. *downstairs*, *inside* vs. *outside*, *come* vs. *go* etc. Subjectivity and objectivity, which hinge on vantage point, concern the degree to which a given element is fully profiled, put onstage, and, thus, objectively construed or relegated to the offstage region, being thus unprofiled and subjectively construed. An increase in subjectivity is illustrated in the construction *going to* in the following sentences: *She's going to a shop* (fully profiled physical motion) vs. *She's going to burst out crying* (no motion, intentionality, increased control by the speaker) vs. *It's gonna rain heavily tonight* (no motion, no intentionality, full control by the conceptualiser). The third factor essential for construal, prominence, is evident in profiling, i.e., the process of zooming in on an entity in the onstage area, or in trajector/landmark identification. Profiling can be illustrated in *Max jutted his elbow into big Sam's ribs* (from Contemporary Corpus of American, hereafter COCA, Davies (2008)), where the element being focused on is *elbow* and it is singled out against the immediate scope of the *arm* and the maximal scope of the human body. The other aspect of prominence, trajector/landmark identification, comes to the fore in sentences such as *The good harness for the horse* (tr) *was behind the table* (lm) (from COCA) vs. *The nice table* (tr) *was in front of the harness for the horse* (lm). Finally, the variable referred to as focusing involves selecting relevant "conceptual content for linguistic presentation" and arranging it in line with what is to be foregrounded and what should remain back-

grounded (Langacker 2013: 57).

As we shall see in the present study, the choice between the prepositions *of* and *about* as alternate complements of communication and cognition predicates is ultimately not just a matter of lexical variation, but of divergent construals of the conceptualised scene. The differences in construal are explored in Section 1.2, but in the most general terms they relate to the scope of conceptual information that the speaker possesses with respect to the object and on which s/he chooses to focus. The preposition *of* is more limited and focuses exclusively on the entity put on-stage, whereas *about* has an extended profile. This contrast is manifested clearly in the following sentences (from Google search engine & BNC (Davies 2004)):

(1) *I want a relationship where everyone knows of us, but knows nothing about us*.

(2) *… when one thinks about using computers in schools, one tends to think of technology*.

Example (1) makes it evident that *knowing of something* involves mere awareness, whereas *knowing about* implies being in possession of much more sensitive information. Similarly, in (2) it is clear that *technology* is the general background against which one thinks about *using computers*. Both *thinking* and *knowing about* entails considerably more extensive processes, here mental processes, performed on the objects and a broader scope of conceptual content evoked in the conceptualization. Let us now look in more detail at what construal differences can be identified for the two alternate prepositional complements of verbs such as *think* or *know*.

1.2. *The prepositional complements* of *and* about

To use Rudzka-Ostyn's (2003: 180–183) cogent phrasing, the preposition "*about* is dispersion", be it in the domain of physical, conceptual, or communicative space. It is, therefore, marked by a certain degree of indeterminacy (Dirven & Radden 2007: 329) and imprecision (O'Dowd 1998: 65), which can be understood in terms of literal or metaphorical movement "in any possible direction" (Radden 1981, cited in Dirven 2003: xvii; Dirven et al. 1982: 58, 79). When used figuratively, in the abstract spheres of thought or speech, *about* can be interpreted to concern "mental motion on topics" (Rudzka-Ostyn 2003: 181). As noted by Dirven et al. (1982: 27, 57) or Lindstromberg (2010: 141), this preposition is, in fact, the standard indicator of a given object being the topic. *About* as a prepositional complement of verbal and adjectival expressions (e.g., *crazy about*, *argue about*) is likely to have once

been associated with the speaker's "attention enveloping the Landmark" (Lindstromberg 2010: 255). When combined with a cognition predicate such as *think* or a communication verb such as *talk*, the preposition *about* can be seen as implying that the subject's interest encompasses not only the object itself, but also anything that is related to it and is of relevance (Lindstromberg 2010: 207). In other words, *about*, when accompanying cognition or communication predicates such as *talk*, implies that "all possible aspects of the topic" are considered and so the speaker takes into account a "wider scope" of the issue (Dirven et al. 1982: 60, 62). There are also some semantic and syntactic properties of *about* that support the above observations. As indicated by Dirven et al. (1982: 28–29), in the context of *speak*, *about* makes it possible for the verb to be modified by such adverbs as *in detail* or *for a long time*, which seem highly unlikely with such complements as *of*. The preposition *about* is also associated with clefting and fronting, which, again, is not encountered with *of* (Dirven et al. 1982: 28). Such usage features, identified in a corpus-based study, highlight the fact that the speaker, when selecting the preposition *about* as a verbal complement, focuses on a broader context of the topic under consideration. This can be illustrated in sentence (3) below (COCA), in which it is clearly intimated that the person in question is well versed in addiction:

(3) *I think she did not want to use the word " addict. " She said " weakness, " but she used the language of somebody who knows about addiction and has studied addiction, …*

It can thus be posited that the speaker's knowledge concerning the landmark following the preposition *about* is extensive, which is why Radden (1981, cited in Dirven 2003: xvii) classifies this preposition as "the prototypical" instantiation of his "general notion of Area".[2] The above-mentioned properties of generality, indeterminacy, or imprecision inherent in the profile of *about* affect the character of the relation holding between the verb phrase and its object, which is likely to be extended, dispersed, and unfocused in nature.

The other preposition examined in the present study is *of*. As Langacker (1991: 37; 1999: 74) indicates, *of* denotes an "intrinsic relationship between its trajector and landmark", with the former being "an inherent subpart" of the latter. The intrinsic character of this "restricted-subpart relationship" is additionally accentuated by the phonological reduction and cliticization of the preposition to the point where

[2] Area, also called "theme", concerns "the frame [within which] … certain actions, states, or events are located" (Radden 1978: 328, cited in Dirven 2003: xvii). It is juxtaposed with *Patient* and the difference is illustrated by Radden (1978) in examples such as *He told me his secret* (Patient) vs. *He told me about his secret* (Area).

it becomes hardly audible (Langacker 1999: 77). Its formal and semantic tenuousness reinforces the inherence between the "relational participants" by minimizing "the conceptual distance between" them (Langacker 1999: 77). To illustrate this, we can consider sentences such as (4) below (from COCA), where, clearly, the focus lies exclusively on the element expressed verbally:

(4) *Do you know of specific individuals who are in charge of those terrorist organizations?*

To the same effect, Lindstromberg (2010: 206–207) states that *of* has an "integrative meaning", which means that the object of the preposition is integrated into the event. This can be illustrated by such pairs of sentences as *I think of Socrates* vs. *I think about Socrates* (Lindstromberg 2010: 207). When *think* is complemented by *about*, the object of thought extends beyond its immediate referent to also incorporate any of its aspects that might be of importance to the subject (Dirven et al. 1982: 26, 60–62; Lindstromberg 2010: 207). In addition, not only is the object of thought extended conceptually, but also the process of thinking may extend temporally over a longer period (Dirven et al. 1982: 29; Lindstromberg 2010: 207). On the other hand, when the mental predicate is followed by *of*, the speaker concentrates primarily on the object itself and the process of focussing attention may be temporary and short-lived (Dirven et al. 1982: 27; Lindstromberg 2010: 207). Hence, predicates complemented by *of* are highly unlikely to collocate with adverbial expressions such as *for a long time* or *in detail* (Dirven et al. 1982: 28–29). In a similar vein, Taylor (2002: 196–197), working in the framework of Cognitive Grammar, specifies that the intrinsic relation conveyed by *of* obtains between "a profiled entity and an entity in the base". In our case, the base can be perceived as the relevant domain against the background of which the object is highlighted and zoomed in on. It is interesting to note that *of*, in itself, is rather semantically impoverished, which is not to say that it is meaningless (Langacker 1999: 73). What it means is that it cannot be perceived as genuinely polysemous (Taylor 2002: 325). It has merely a "schematic value, which unifies with the more contentful relation present in the semantic structure of the trajector noun" (Taylor 2002: 325). Drawing on this aspect of the semantic haziness of the preposition *of*, Radden (1981, cited in Dirven 2003: xviii) identifies it as instantiating an "origin Area", which he contrasts with the general Area introduced by *about*. In this respect, unlike *about*, the preposition *of*, when preceding the object of mental or communication verbs, indicates that the speaker has only "limited information" at his or her disposal. It also implies that only relevant aspects of the topic are considered and so the focusing of attention is very selective (Dirven et al. 1982: 79).

The above observations lead us to formulate two hypotheses with regard to the usage patterns that are expected to be associated with the two prepositions. It is posited here that the choice between *of* and *about*, rather than being purely a matter of lexical variation, is a choice between two alternate construals of the same scene. The preposition *about*, which is characterized by indeterminacy and extended scope, will be more prone to be correlated with abstract objects. This is because such objects are more intangible, more likely to require reference to some additional specifying information, and are therefore more disposed to exhibit such features as imprecision or dispersion, which are typical of *about*. *Of*, on the other hand, distinguished by the intrinsic and focused nature of the relation it establishes between the two participants it integrates, is hypothesised to be more attracted to concrete objects, which are more easily graspable and, therefore, can be zoomed in on without any difficulty. Naturally, it must be stressed that any conceptual content can be construed in either a more general, extended manner or more precisely, which is why the assumptions propounded here should be treated as tendencies, which will be tested statistically. In other words, any semantic object can be construed in both ways, but it is reasonable to suppose that the inherent nature of the referent will weigh upon the speaker's construal choice. This would suggest that the object, by its very nature, will be more amenable to either the extended perspective or the nuclear focusing of attention.

2. Method and data

2.1 *Spoken British and American English*

At an analytical level, the present study focuses on a constructional alternation. The application of quantitative usage-based methods to such lexico-syntactic alternations has an established tradition (Gries 2003; Heylen 2005; Grondelaers et al. 2007; Bresnan et al. 2007; Szmrecsanyi 2010 *inter alia*).

In the present context, the constructions under analysis can be schematically represented by VERB + *of vs*. VERB + *about*. The verbal slot has been restricted to two types of predicates, namely, communication verbs and cognition verbs, limited to *speak*, *talk*, *know*, and *think*. These two classes of predicates are semantically divergent, with one group designating the innermost subjective mental processes and the other standing for external intersubjective interactive processes. However, despite this inherent difference, they both report on the conceptual dimension, as *construed* by the cogniser. It is assumed here that depending on whether the speaker chooses a more extended profile or a narrower scope for the object, he or she will

select *about* or *of*, respectively. The choice of such two disparate verb classes is designed to help us identify the contribution of the predicate semantics to the constructional construal. More specifically, it is intended to establish whether the patterns of use to be revealed are determined by the lexical semantics of the predicates or the construal characteristics imposed by the prepositions. What this means is that if we find that distinctive features are associated with *of* or *about* and that they systematically occur across the four verbs and the two verb classes then we can be reasonably sure that these associations are part of the constructional construal rather than predicate semantics.

The data in this study were extracted from the spoken components of the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). The total number of occurrences of the constructions subjected to analysis is nearly 700. The proportions relative to the construction and dialect are specified in Table 1.

Table 1. Proportional sample size of prepositional complement

about	Dialect		*of*	Dialect		
predicate	UK	US	predicate	UK	US	total
know	30	30	*know*	30	40	130
speak	51	56	*speak*	51	51	209
talk	44	42	*talk*	44	41	171
think	40	45	*think*	42	41	168
total	165	173	total	167	173	678

The examples were manually annotated for a wide range of formal and semantic features, which we will enumerate and discuss in Section 2.2. Following the qualitative analysis, the data were submitted to exploratory and confirmatory statistical modelling with a view to testing the hypotheses (cf. Section 1.2) and revealing the behavioural profiles of the constructions.

2.2. *Multifactorial usage-feature analysis*

The principle of usage-feature analysis was developed independently by Dirven et al. (1982), Rudzka-Ostyn (1989), and Geeraerts et al. (1994) in Cognitive Linguistics and by Atkins (1987) and Hanks (1996) in corpus linguistics. The application of multivariate statistics to the results of usage-feature analysis was pioneered by Geeraerts et al. (1999) and Gries (2003) and later developed into a fully-fledged methodology within Cognitive Linguistics (e.g. Gries & Stefanowitsch

2006; Glynn 2009; Gries & Divjak 2009; Glynn & Fischer 2010; Divjak 2010; Glynn & Robinson 2014 *inter alia*). The principle lines of research in the field have focused on near-synonymy and contructional alternations. However, more recently, there have been attepmts at extending the method to highly subjective and abstract semantic features typical of theoretical research in Cognitive Grammar and discourse analsyis (cf. Glynn 2010a; 2016a; 2016b; Krawczak 2014; 2015; 2018; Krawczak & Glynn 2015). Operationalising abstract semantic characterstics of use remains a fundamental challenge for the application of this method to cognitive linguistic research questions.

The principle of multifactorial usage-feature analysis itself is straightforward and permits the quantification of traditional techniques of linguistic analysis. Large numbers of whatever linguistic phenomenon is under investigation are extracted from corpora. These examples are then analysed, manually and systematically, for a range of usage-features that pertain to the research question. Such features can include any formal, socio-contextual, or semantic characteristic of language. The resulting usage-feature annotation consists of large detailed multifactorial profiles of the phenomenon under investigation. In order to identify patterns in the usage-profile (patterns argued to represent language structure), multivariate statistics is employed. The use of the statistics not only permits the identification of the patterns but also calculates the likelihood that such patterns would be found again in re-sampling. It also determines the descriptive power of those patterns by estimating how much of the behaviour of the data the patterns can predict. For further discussion of the strengths and weaknesses of the method cf. Glynn (2010b; 2014a).

The constructions under analysis were manually annotated for a set of usage features chosen to operationalise the distinction between the construals but also to account for other semantic and formal features that may play a role in the linguistic configuration of the construal. These features are listed in Table 2.

Table 2. Annotation schema of formal and semantic usage features

Observable Factors	Features	Subjective Factors	Features
Subject form	noun, pronoun, proper noun	Subject semantics	human, non-human
Subject person	1st pers., 2nd pers., 3rd pers.	Modal semantics	ability, possibility
Subject number	singular, plural		obligation, hypothesis
Object length	short, long	Adverb semantics	addition, epistemic, intensifier
Object form	noun phrase, pronoun,		location, manner, temporal
	proper noun, gerund, clause	Object semantics	abstract, activity, event, human
Adverb modification	adverb, no adverb		object, place, state,
Sentence temporality	present, past, future		state of affairs
Sentence polarity	positive, negative	Speaker engagement	marked, unmarked
Epistemic modality	indicative, conditional	Speaker evaluation	negative, neutral, weak
Propositional modality	interrogative, declarative	Topic of discourse	social, private

We will here only exemplify these features that proved significant in their contribution to the structuring of the data, leaving the others out of the discussion. The most crucial factor is that of object semantics, which is here used to operationalise the hypothesised difference in construal imposed by *of* vis-à-vis *about*. As indicated at the end of Section 1.2, it is assumed here that *about*, whose profile is intrinsically indeterminate and imprecise, will tend to be associated with more abstract objects, which are similarly indefinite and nebulous. If we look at examples (5), (7), (11), (12), we can see that the relational objects in them (*his ordeal*, *the opening*, *independence*, and *paying more attention*) are far from nuclear notions. Rather, they require a much broader contextualization and specification so as to be clear. For that reason, it is more natural to be *thinking* or *talking about* them, as one makes conceptual reference to other relevant categories that are needed for the concepts to be actually meaningful. It is also more probable that such abstract objects will figure in the subject's thoughts for longer and will be discussed more thoroughly. The preposition *of*, on the other hand, which minimizes the distance between the two participants of the construction and which zooms in specifically on the object expressed without being sidetracked, is expected to be attracted to more concrete objects. Such objects are more readily graspable and can be easily zoomed in on (consider examples (6), (8), (9), (10)). For instance, in sentence (9), *weapons* constitute a concrete thing that is the exclusive focus of attention, thus being more likely to be related to the predicate by means of *of*. Naturally, these hypotheses concern tendencies, rather than absolute rules and they will be tested in statistical models. The sentences below (from the dataset) illustrate the usage features identified for object semantics:

(5) *One of the former pupils, who's now nineteen, has been speaking about his ordeal*. (abstract)

(6) *We don't trust the Zionist enemy at all, because he only thinks of killing more Palestinians. But we trust the Palestinian resistance and its leaders.* (activity)

(7) *He hadn't seen any posters in the village. Would people know about the opening*? (event)

(8) *But between the tears and the crushing sorrow, Elizabeth thinks of Charlie and Braden*. (human)

(9) *I don't know of any other weapons, at this point, that we're trying to do a trace on*. (object)

(10) *… she speaks of the new South Africa*. (place)

(11) *As a private person, as man who was born in another small Republic of the Soviet Union Georgia which also thinks about independence I sympathize with Lithuanians and understand their drive toward independence*. (state)

(12) *Bill Clinton pays more attention to Hillary than to you. What do you think about that, Mr. Vice President*? (state of affairs)

Other factors that turned out to play a role in structuring the data are as follows: subject person, object length, topic of discourse, evaluation, and engagement. Let us discuss them more closely and illustrate with specific examples. The first variable is self-explanatory, distinguishing between the first, second, and third grammatical person of the subject. With respect to the factor of object length, it was determined on the basis of the number of words of the object so that anything of more than 5 words was regarded as a long object. Topic of discourse is further subdivided into social, as in (13) and private, exemplified in (14). It is expected that social topics, which are more remote and of which the speaker has less knowledge, will be more likely to be associated with *of*, as in example (13). While in the case of personal matters, where the subject is fully aware of all the intricacies and in which he / she is immersed, the extended construal imposed by *about* is more probable, as illustrated in (14).

(13) *Did you find any evidence on this matter you've discussed, or any other matter, that the President of the United States knew of these reported illegalities?*

(14) *He very seldom spoke about his family.*

The factor of evaluation subsumes three features: positive, negative, and neutral, exemplified respectively by sentences (15), (16), and (17) below. It might be hypothesised that when the evaluation is negative or positive, the context of the object will be extended, which is why *about* could be more commonly a complement of the predicate in such cases. This is because intense emotions are prone to be invoked by conceptually more complex situations associated with entities that will be similarly complex.

(15) *When she speaks about energy, she sounds brilliant.*

(16) *the Government troops which were sent in to counter the attack at Buto ... spoke about indiscriminate killings, rapes and about looting and burning*

(17) *Firstly, when one thinks about using computers in schools, one tends to think of technology*

In example (15), it is clear that the speaker to be impressed by the brilliance of the woman must have had considerable exposure to her discussing the issue. Likewise, in (16), the monstrosity of the events is certain to have been described in detail by the government and the speaker thus conveys the large-scale character of the events.

Finally, engagement can be either marked, in which case it will typically also be strongly evaluative, or unmarked. In the former case, it is assumed that the speaker is more likely to construe an object that is highly engaging in an extended manner.

(18) *Do you know of any shops selling ski wear which looks good and is reasonably priced*?

(19) *Jackson said Obama was " talking down to Black people " when he speaks about the need for Black fathers to be more responsible for their families. " I wan na cut his nuts off, " Jackson said.*

Example (19), which is also an instance of strongly negative evaluation, shows that the speaker chooses this extended construal for a topic which is particularly important to him/her and which he / she must have witnessed amply. The neutral examples, for both evaluation and engagement, on the other hand, are more specifically oriented on the objects in question (*technology* and *shops*). This focus of attention, however, is rather superficial in the sense that they serve as a background for conceptualizing something else (*using computers* and *ski wear*).

3. Results and discussion

3.1. *Predicate or prepositional construal*

The first step is to isolate the factors that are interacting exclusively with the prepositional complements. In order to do this, we need to identify which factors, if any, correlate with the prepositional uses, regardless of the predicate class or individual predicate involved. One simple way of identifying such distinctive patterns is to cluster the predicates with various sets of factors until a clustering is found which groups the examples relative to the complement and not the predicate.

Hierarchical Agglomerative Cluster analysis works by considering the usage-feature profile of each occurrence and then systematically grouping each occurrence with another occurrence to which its profile is the most similar. If a set of factors produces a usage profile that results in the examples being grouped in such a way that the *of* examples are found to be similar to each other and distinct from the *about* examples, then we can suppose that whatever factor was used to generate the profiles of the occurrences is one that interacts with the prepositions, and not the predicate.

Figure 1 presents the results of an agglomerative hierarchical cluster analysis of the examples using the object semantics.[3] The examples are categorised as combinations of the predicate and the prepositional complement. The results are clear and indicate that the object semantics is interacting with the predicate semantics and not the prepositional complementation. The numbers below the clusters are the sequential order in which the clustering algorithm allocated the splits in the dendrogram. The numbers above, on either side of the branches, are bootstrapped confidence scores. The number to the left (au) is an unbiased probability, calculated with multiscale bootstrap resampling and the number to the right (bp) is a standard bootstrap probability. The former is argued to be more accurate (Shimodaira 2004). The bootstrapped confidence scores are all relatively high.

Two distinct high-level clusters are revealed, distinguishing between communication predicates and cognition predicates. These two high-level clusters are, in turn divided into two subclusters, based on the individual predicates. In this analysis, the complement appears to have no bearing at all on the clustering. In other words, using the object semantics to 'sort' the examples finds a clear lexical semantic map of the near-synonymy of the predicates, regardless of the complement used in the example.

[3] Analysis performed with the R package pvclust (Suzuki & Shimodaira 2011). For an explanation of cluster analysis, cf. Divjak & Fieller (2014).

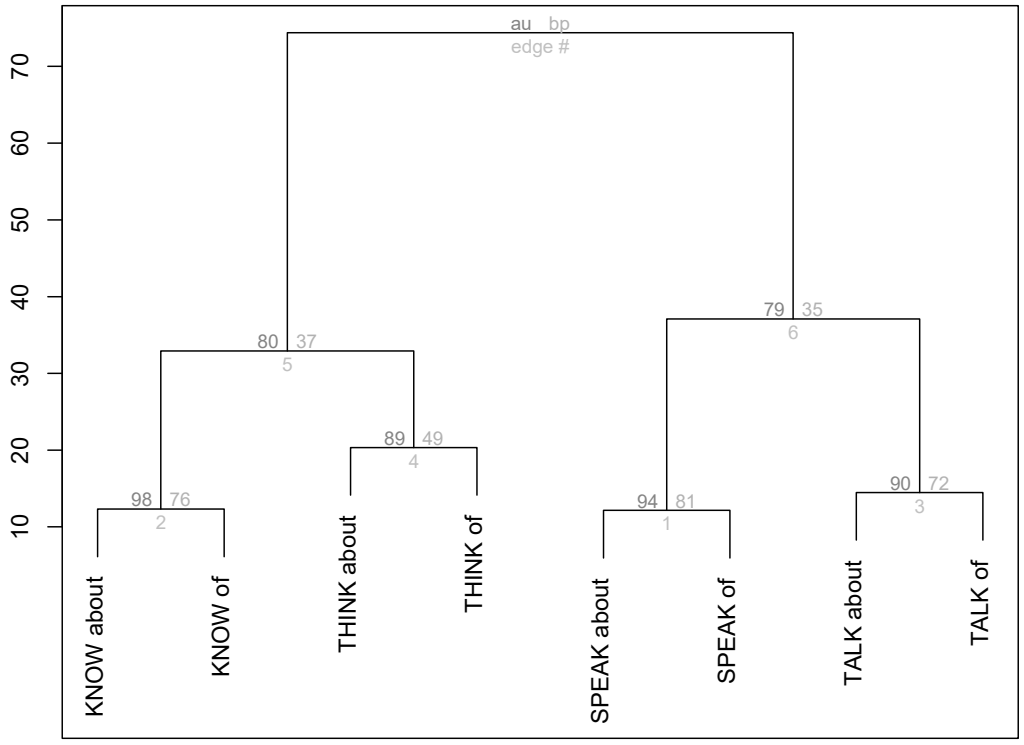

Figure 1. Complement-predicate combination clustered by direct / indirect object semantics Hierarchical cluster analysis (dist. matrix: Euclidean, agglomeration method: Ward)

The clustering was repeated for every combination of semantic and formal features including the subject semantics, the utterance pragmatics, and morpho-syntax in various combinations. No combination of features resulted in a clustering of examples relative to prepositional complement. This demonstrates that whatever the conceptual construal of the complement, the predicate semantic profiling is more pronounced. Although it may be obvious that the semantic contribution of the predicate is greater than its complementation, the systematicity with which the predicates are correlated with the object and subject semantics is striking.

A second step in order to determine whether we can isolate the semantics associated with the complement is to run Pearson's Chi-square test for independence on the two sets of examples (*of* and *about*).[4] The test reveals statistically significant independence between the two sets of examples, but only at an alpha level of 0.05

[4] Cf. Gries (2014) for an explanation of the Chi-square test and the interpretation of the residuals.

(df = 7, *p*-value = 0.01002). Moreover, residuals of the test show that only two of the object semantics features play any important role in distinguishing the two complement construals. Only 'state of affairs' (SoA) and 'place' reveal any substantial effect size in distinguishing the two complements. However, that 'state of affairs' is associated with *about* and 'place' with *of*, does support the hypothesis that the former will be used to profile more abstract objects and the latter more concrete concepts. Table 3 includes the Pearson residuals for the chi-square test.

Table 3. Pearson's residuals for the Chi-Square test *of* vs. *about*

	Object semantics							
Compl.	Abstract	Activity	Event	Human	Object	Place	State	SoA
about	0.4419186	-0.03032530	-0.8170300	-0.8794487	0.3035067	-1.414484	-0.3335172	2.302253
of	-0.4464314	0.03063497	0.8253733	0.8884294	-0.3066061	1.428928	0.3369230	-2.325763

Next, in an attempt to control for the role of the predicate semantics and isolate the correlation between the object semantics and the complement, the data were submitted to a binary mixed effects logistic regression. [5] Table 4, below, summaries two mixed effects logistic regression models. These models are designed to predict the prepositional complement, of the four predicates, as *of* or *about*. The four predicates (*think*, *know*, *speak*, *talk*) and the two dialects (British and American) are included as the random variables in the model. This means that any impact the different predicates or the dialect variation has on predicting the outcome (*of* vs. *about*) is accounted for and excluded from the results.

The models in Table 4 were run using the Laplace approximation algorithm. The *p*-values of the estimates were checked with a model using maximum quasi-likelihood estimation (logit) and Markov chain Monte Carlo methods and no noteworthy differences were found. The significance codes for the alpha levels in Tables 4 are: *** < 0.001, ** < 0.01, * < 0.05, . < 0.1

[5] The mixed effect logistic regression was performed in R, using the lme4 package (Bates & Sarkar 2007) and LangaugeR package (Baayen 2008). For an explanation of mixed effects logistic regression cf. Faraway (2016).

Table 4. Mixed effects binary logistic regression (Laplace estimation)

Coefficients	Model 1 Estimates	*p*	Model 2 Estimates	*p*
Object Semantics - HUMAN	0.45396	.	0.32811	
Object Semantics - EVENT	0.44498	.	0.31340	
Object Semantics - ACTIVITY	0.23779		0.08515	
Object Semantics - OBJECT	0.04847		-0.07736	
Object Semantics - PLACE	1.18616	*	1.00145	*
Object Semantics - SoA	-1.18173	**	-1.16760	**
Object Semantics - STATE	0.19223		0.26747	
Topic of Discourse - Social	0.44383	*	–	
Subject Person - 1st	0.68088	*	–	
Subject Person - 3rd	0.73352	*	–	
Evaluation - Negative	0.70530	**	–	
Evaluation - Positive	0.82672	**	–	
Engagement - Marked	-0.62301	**	–	
Diagnostics				
VIF (maximum)	3.334		1.072	
κ	8.609		2.533	
AIC	940.9		955.8	
BIC	1013		1001	
Cox Snell R^2	0.065		0.028	
Nagelkerke R^2	0.086		0.037	
ROC	0.637		0.571	
Somer's D_{xy}	0.274		0.143	

The models were checked for outliers and influential observations and extensive diagnostics were performed.[6] Particular care was taken to ensure model orthogonality since many of the predictor variables could have produced problems of collinearity. For this reason, two tests of multicollinearity were performed and the maximum level of collinearity identified for each model is listed. The methods used are: (i) the variance inflation factors (VIF), which should be beneath 0.4, and (ii) a Kappa statistic (κ), where <10 is acceptable <30 is moderate, and >30 represents problematic collinearity.[7] The models were obtained using a backward selection

[6] Specifically, the diagnostic recommendations outlined in Agresti (2002), Gellman & Hill (2007), Baayen (2008), Thompson (2009), and Faraway (2016) were followed.

[7] For a discussion on the maximal VIF, cf. Glynn (2010c; 2014c) and Speelman (2014). For a discussion on the rules of thumb using a Kappa statistic to measure collinearity, cf. Gorman (2009) and Dormann et al. (2012).

procedure, based on parsimony and the Akaike information (AIC) and Bayesian information (BIC) criteria. Model 1 is the maximally parsimonious model and model 2 the simplest. The pseudo R^2s, the Sommer's D_{xy} and C(AUC) model statistics, for both models, are low which indicates that neither model is a good fit. This overdispersion does not mean that we cannot interpret the table of coefficients, but we must take extreme care in any conclusions thus reached.

To interpret the coefficients in Table 4, positive estimates predict an *of* complement and negative estimates an *about* complement. Although, Model 1 offers a range of significant predictors, how they contribute to the falsification or the support of the hypothesis is not clear. Tentative interpretation can, however, be made. The preposition *about* is here attracted significantly to abstract 'states of affairs'. As postulated in Section 1.2, given the extended scope of predication implied by this complement, it is more likely that it will be associated with abstract objects. This is because such objects are characterized by indeterminacy and imprecision, which are typical of the construal imposed by *about*. *About* is also predicted by 'marked' engagement. It is intuitively interpretable that topics that invoke intense emotions and engage the subject strongly are prone to be discussed or considered at length. Hence, the broader perspective introduced by *about* is more likely. The other preposition *of* is predicted by object semantics designating 'places', which are concrete and so more easily focused upon. This supports our hypothesis. *Of* also correlates, but less distinctly, with positive and negative evaluation, social topics, and first as well as third person of the subject. These findings are less obvious to interpret. We would have actually expected positive or negative evaluations to predict for *about* for the same reason as mentioned above for engagement. The association of *of* with social topics can be accounted for in that the speaker normally possesses less knowledge about such topics than about more personal matters, where the knowledge is extensive. Finally, the association between *of* and the first and third person correspond to a parallel "curious finding" revealed in the corpus study conducted by Dirven et al. (1982: 29), where the non-first-person perspective was significantly linked to the *of* complement.

However, since we have no evidence that such indirect interpretation informs the abstract – concrete hypothesis, it is safer to restrict the interpretation to Model 2. During model selection, subject semantics and topic of discourse both revealed significant predictors, but none as important as those found in object semantics. When combined with subject semantics, these factors either revealed collinearity or were no longer significant. Interestingly, neither length of the object nor its type showed any significant prediction for either of the complement. This in itself is a surprising and informative result. It shows that, despite the seemingly obvious dif-

ference between the two construals, in actual use, they are formally (at least), extremely similar.

Turning to the direct object semantics. The model confirms what the chi-square test revealed, but the addition of the random factors in the model assures us that the difference is not an epiphenomenal effect of the predicate or even dialect variation.

3.2. *Constructional construal*

The reason for the lack of clear results can be accounted for in four, non-mutually exclusive, ways:

(i) There is little difference between the construal of the two prepositional complements;

(ii) The operationalisation of the semantic profile of the object does not capture the construal differences;

(iii) The lexical semantics have a stronger impact upon the object semantics than the prepositional complement, in effect, hiding its effects on the variation in the data;

(iv) The predicate – prepositional complement pair has a specific constructional construal that is non-predictable from its composite parts.

The first possibility is not only counter to the form-variation – meaning-variation principle of Cognitive Linguistics, it goes against simple intuition. The second possibility can only be determined by re-analysis. However, with further multivariate investigation of the results of this analysis, it may be possible to tease apart the object semantic correlation with the predicate and prepositional semantics and / or identify constructional pairings.

A first exploratory investigation, in the form of a multiple correspondence analysis, reveals some prepositional complement systematicity. Figure 2, below, presents the biplot of a correspondence analysis that examines the simultaneous associations between the prepositions, the predicates, and the object semantics. The size of the 'bubble' assigned to the data points indicates the contribution to the squared cosine (cos^2) and proximity between data points represents degree of association. The higher the cos^2, the more important the given data point is in explaining the behaviour of the data.

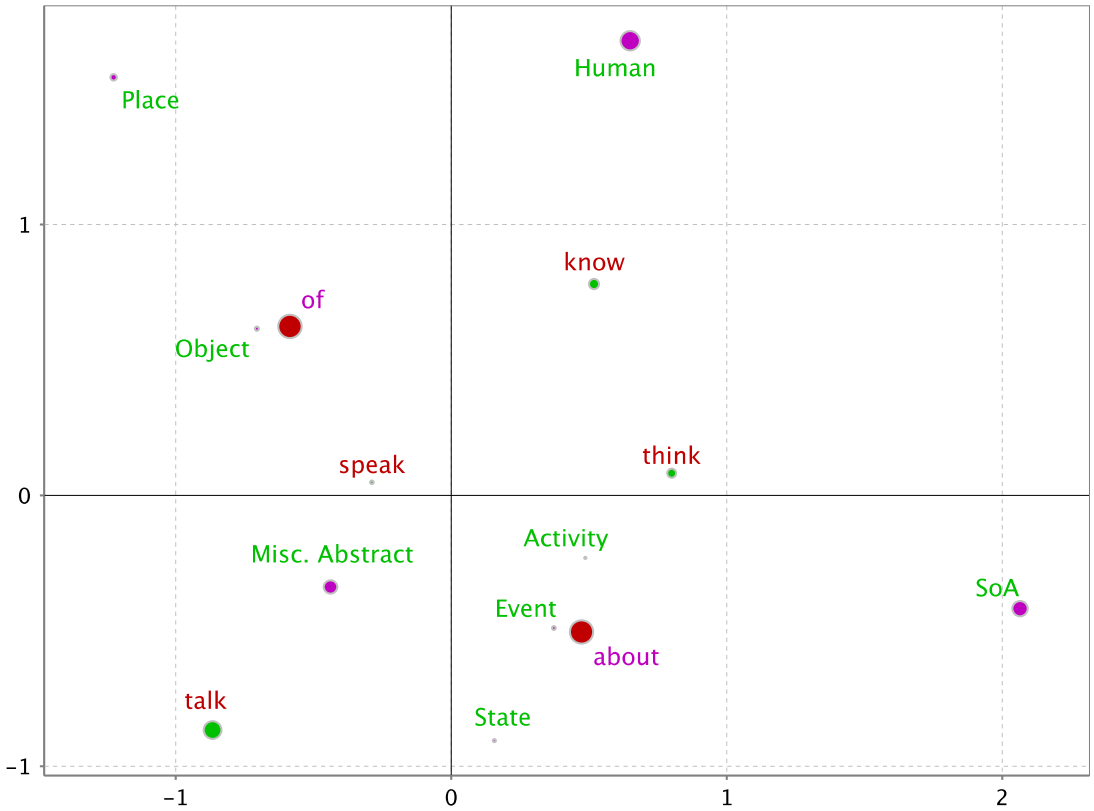

Figure 2. Association of complement, predicate and object semantics
Multiple correspondence analysis (Burt matrix method)[8]

Proximity between the data points is entirely relative, so that a data point, relatively far from another point, but still closer to it than it is to other data points, indicates distinct association. The Burt matrix algorithm used in the analysis in Figure 2, does not produce an interpretable score of explained variance (inertia). Therefore, before we interpret the analysis, we need to examine some of the numerical output of two other correspondence analyses using different algorithms. Table 5 summarises the so-called explained variance (inertia) of the analysis using Greenacre's (2007) so-called joint method of multiple correspondence analysis.

[8] Note that the correspondence analysis was performed in both the FactoMineR (Husson et al. 2012) and ca (Nenadic & Greenacre 2007) packages. The numerical summaries, quality scores and scree plot were produced using ca and Greenacre's (2007) "adjusted" method, where the biplot was produced using a standard Burt matrix and the FactoMineR package. There was no noticeable difference in the plots produced by the Burt and 'adjusted' correspondence analyses. The FactoMineR package was used for the biplot because of its superior graphics options.

Table 5. Explained variation, multiple correspondence analysis (Greenacre adjusted) Complement, predicate, and object semantics[9]

Principal inertias (eigenvalues):		
Dim	**% Explained of structure (cumulative)**	**Scree plot**
1.	44.4	**********
2.	15.4 (59.8)	*********
3.	12.1 (71.9)	*******
4	3.1 (75.0)	**

Applying this algorithm to the analysis, we obtain an explained inertia of 59.8%. The same analysis performed with Greenacre's adjusted algorithm results in 79.3% explained inertia. Although these scores are relatively low in terms of stability, they still suggest interpretable results. The greatest concern is that the low score in the adjusted method is a result of the fact that the analysis requires 3 dimensions to properly represent the structure of the data. This is indicated by the scree plot and the list of contributions in Table 5.[10] Although the two-dimensional plot is a relatively reliable map of the associations between the different forms and the object semantics, the two-dimensional visualisation is missing an important part of the structuring. The first dimension, visualised along the *x*-axis of the plot, accounts for 44.4% of the structure, the second demission along the *y*-axis accounts for 15.4%, but a third dimension which is not depicted would account for another 12%. Although 3-dimensinal plots are possible, just as it is possible to present three two-dimensional plots with different combinations of the three dimensions, such visualisations are difficult to interpret due to their complexity.

In Table 6, the contribution to the first two dimensions of each of the features is listed as well as the so-called "quality" score, which is a calculation of the relative reliability of the depiction. It is here that we see the implications of the relatively poorly explained inertia. As a rule of thumb, reliably depicted data points should have a quality score of at least 500 (Greenacre 2007). Although overall, the scores are reasonable, note that the two data points in which we are most interested both rate at 419. It seems that although the plot is interpretable, there are interactions between the complement, the predicate, and the object semantics that are not being identified.

[9] Scree plot and table of contributions (Table 5) were produced using R package ca (Nenadic & Greenacre 2007).

[10] For an explanation of the interpretation of correspondence analysis, cf. Glynn (2014b).

Table 6. Quality and contribution, multiple correspondence analysis (Greenacre adjusted) Complement, predicate, and object semantics

Feature	Greenacre Quality score	Contribution dimension 1 (*x*-axis)	Contribution dimension 1 (*y*-axis)
Preposition *about*	419	15	83
Preposition *of*	419	15	84
Predicate *know*	678	167	79
Predicate *speak*	596	95	47
Predicate *talk*	537	112	3
Predicate *think*	708	97	197
Object Semantics 'Activity'	561	8	322
Object Semantics 'Event'	719	1	203
Object Semantics 'Human'	469	102	9
Object Semantics 'Misc. Abstract'	673	72	34
Object Semantics 'Object'	570	10	56
Object Semantics 'Place'	'80	1	11
Object Semantics 'SoA'	703	296	59
Object Semantics 'State'	514	9	22

Turning to the actual results of the correspondence analysis, calculated using the Burt Matrix, we see clear associations between prepositional complements and certain features of object semantics. Although there is no clear clustering, note that the most important contributions to the structuring of the data are the two prepositions and the object semantics 'human'. This is represented by the size of the actual data point; larger meaning higher $\cos^2$. This finding is crucial since it means that although there is instability in the results and complexity that the analysis is not able to account for, we know that when we separate out all the factors, predicate, preposition and object semantics, the role of the preposition, and therefore the construal, is important.

The association between the preposition *of* and the object semantics 'place' is strong and with 'object' it is strong and distinctive, but specifically for the predicate *speak*. The preposition *about* appears to have a less distinctive association except for the object semantics 'state of affairs'. It also reveals non-distinctive association with 'state', 'event' and 'activity'. This is all relative to the three other predicates, which appear stretched across the two quadrants with which *about* correlates. In the top right quadrant, the lexemes *know* and *think* cluster vaguely as a semantic class, distinctly associated with 'human'. In the bottom left quadrant, the predicate *talk* is associated with 'miscellaneous abstract'. These two sets of relations explain why the previous analyses have not been able to capture the propositionally encod-

ed construal. The patterning seen here in Figure 2 shows that two of the semantic features, 'human' and 'miscellaneous abstract', are important to structuring the behaviour of the data, but that they are correlated with specific verbs more than either of the prepositions. Nevertheless, the plot reveals, once again, evidence in support of the hypothesis that more concrete objects are more likely to be profiled with the *of* construal and less concrete objects with *about*. In the plot, we see clear and distinct correlations between 'place' and *of* and between 'state of affairs' and *about* that have been revealed in previous analyses, but we also see subtler interactions between 'object' and *of* and between 'state', 'event' and *about*. Both these tendencies are exactly in line with the hypothesis.

Although the correspondence analysis has given us some insights into how the prepositional complements might be construing the conceptualised scene, the picture is not clear and there are reliability concerns with the two-dimensional visualisation. This is likely to be the result of three-way interactions between the complement and predicate relative to the object semantics. In order to determine the reliability of the conclusions drawn from the results of the correspondence analysis, we can turn to a loglinear analysis. Unlike correspondence analysis, this is a confirmatory modelling technique that determines which associations or disassociations are statistically significant.

Figure 3 presents a mosaic plot visualisation of the significance and effect size of associations in the loglinear analysis. The three dimensions of the data, the prepositional complement, the predicate, and the object semantics represent three sides of the plot. Object semantics is positioned at the top with the concrete semantic types to the left and the abstract types to the right. The blue represents association and the red disassociations. If the hypothesis were to be borne out perfectly, we would have a cube with four squares: top left – red, top right – blue, bottom left – blue, bottom right – red. Although the results are far from so clear-cut, such a pattern does emerge to a certain extent. The blue blocks represent significant association and the red blocks significant disassociation. In terms of disassociation, only the interaction of the verb *think* is found to be significant – when combined with *of,* it is not used for two of the abstract object semantics categories. In terms of association, *about*, combined with *think*, *know* and *speak*, is significantly correlated with various abstract object semantics. On the other hand, *of*, when combined with *know*, *think*, and *speak*, correlates with various concrete object semantics. These correlations are not merely tendencies but statistically significant associations. Although restricted to specific verbs and specific object semantics categories, the pattern does confirm the hypothesis.

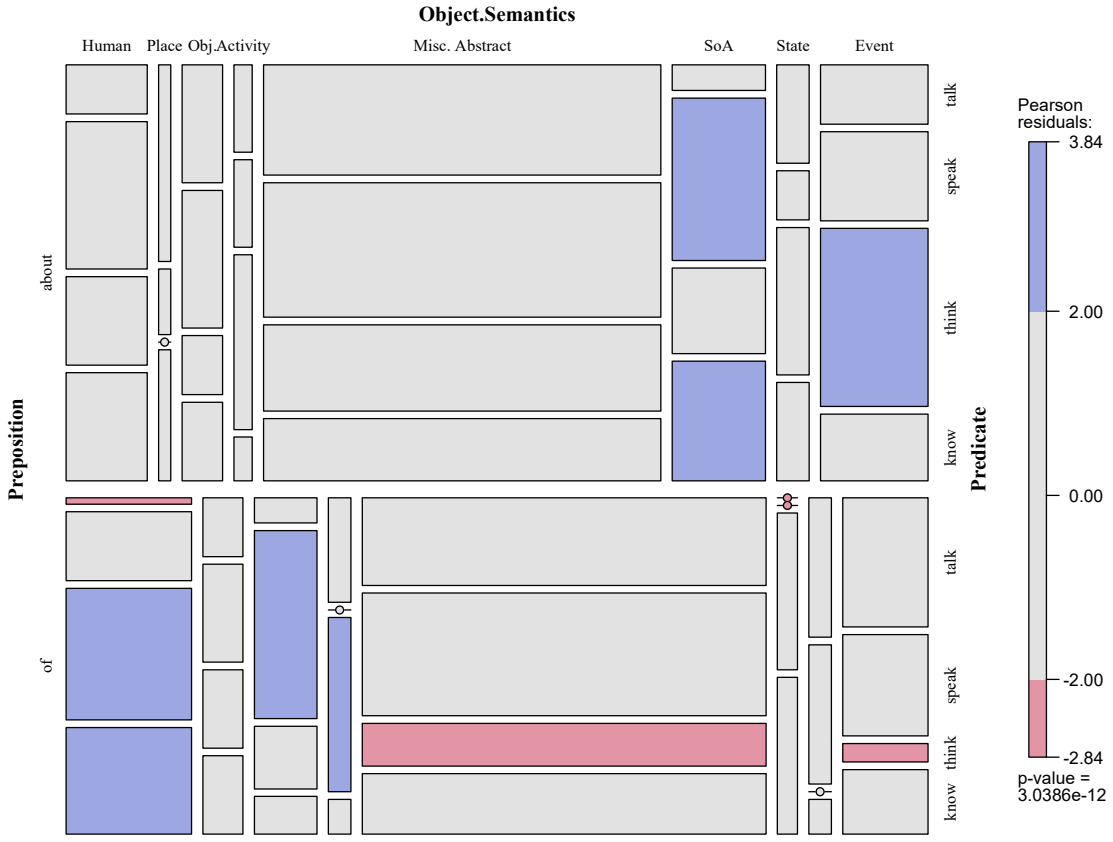

Figure 3. Association model of complement, predicate and object semantics
Results of a loglinear analysis presented in a three-way mosaic plot[11]

It is interesting to note that the verb *talk* does not reveal any significant patterns save a disassociation between *talk*, *of* and 'human' object. This is intuitively sound since *talking of someone* is marked and relatively rare in the corpus. An explanation for this is beyond the purview of the current study, but likely to be due to the specifics of the predicate semantics. Another interesting result that does not directly inform the study is the lack of any significant correlation between *about* and concrete object semantics. It is likely to be the result of another factor, not included in the hypothesis or the operationalisation of the construal. Intuitively, it is reasonable that the VERB *about* construction is less marked than the VERB *of* construction. This idea is corroborated by the raw frequencies of the alternation, which systematically

[11] The loglinear analysis was performed in the R package MASS and the mosaic plot produced with the package vcd (Meyer et al. 2006) in R.

reveal that VERB + *about* is more common than VERB + *of*. Indeed, the frequency difference between the two constructions in the spoken components of both the BNC and COCA is extremely significant (df = 3, p-value < 2.2e-16). In terms of construal, this may suggest that the semantic profile of the *about* prepositional complement is more schematic or less specified than its counterpart. Although the use of the multivariate statistics and the design of this study are not affected by actual differences in frequency, markedness is exactly the kind of phenomenon that does interact with construal. Operationalising such a notion in observational data is a difficult task and one that needs to be treated in future research. However, the results here suggest that it is likely that this factor would need to be integrated into future research in order to properly explain the relationships between the two construals.

4. Conclusion

In order to deal with the problem of not being able to separate predicate and complement semantics, two sets of semantically similar verbs were chosen. It was hoped that by examining the behaviour of *of* and *about*, relative to two different semantic classes, it would be possible to tease apart the semantic contributions. However, neither cluster analysis nor logistic regression, both examining the behaviour of the two complements relative to communication verbs and cognition verbs, systematically revealed structures that can be ascribed to the complements. Instead, what is revealed is that communication predicates correlate with certain object types and cognition verbs with others. The mixed-effects logistic regression analysis, did confirm, however, that significant differences between the two construals, relative to object semantics, do exist especially for the types 'place' and 'abstract state of affairs'. Nevertheless, beyond this, no clear statements can be made about the construal afforded by the prepositional complements.

There could be various reasons for the lack of consistency in the behaviour of the predicates within the two types (communicative and cognition). One such possibility is that construal strategies of different object types varied across the predicates. In order to determine whether this was the case, a multiple correspondence analysis and a subsequent confirmatory modelling in the form a loglinear analysis were performed. This step produced two clear results. Firstly, the operationalisation of the construal alternation has been shown to be sensitive to the individual usage profiles of the predicates. Secondly, although only valid for certain predicates with regard to certain object semantic types, a clear pattern emerged supporting our main hypothesis and no patterns were observed running counter to the hypothesis.

Although the operationalisation of the conceptual relationships involved obviously needs improvement, the proof-of-principle has been obtained and the quantitative falsifiable evidence for the role of the construal in the lexico-grammatical structure has been observed. Future work will need to refine the operationalisation in such a way that the actual profiled difference is integrated into the analysis.

References

Agresti, Alan. 2002. *Categorical data analysis*. New York: Wiley. doi: 10.1002/0471249688

Atkins, Beryl T. Sue. 1987. Semantic ID tags: corpus evidence for dictionary senses. In *Proceedings of the Third Annual Conference of the UW Centre for the New Oxford English Dictionary*, 17–36. Waterloo: University of Waterloo.

Baayen, Harald R. 2008. *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.
http://dx.doi.org/10.1017/CBO9780511801686

Bates, Douglas & Sarkar, Deepayan. 2007. lme4: Linear mixed-effects models using S4 classes, R package version 0.99875- 6.

Bresnan, Joan & Cueni, Anna & Nikitina, Tatiana & Baayen, Harald. 2007. Predicting the dative. In Bouma, Gerlof & Krämer, Irene & Zwarts, Joost (eds.), *Cognitive foundations of interpretation alternation*, 69–94. Amsterdam: Royal Netherlands Academy of Arts and Sciences.

Davies, Mark. 2004-. BYU-BNC. (Based on the British National Corpus from Oxford University Press). Available online at https://corpus.byu.edu/bnc/.

Davies, Mark. 2008-. The Corpus of Contemporary American English: 560 million words, 1990-present. Available online at https://corpus.byu.edu/coca/.

Dirven, René & Goossens, Louis & Putseys, Yvan & Vorlat, Emma (eds.). 1982. *The scene of linguistic action and its perspectivization by speak, talk, say, and tell.* Amsterdam: John Benjamins. http://dx.doi.org/10.1075/pb.iii.6

Dirven, René. 1982. *Talk: Linguistic action perspectivized as discourse*. In Dirven, René & Goossens, Louis & Putseys, Yvan & Vorlat, Emma (eds.), *The scene of linguistic action and its perspectivization by speak, talk, say, and tell*, 37–85. Amsterdam: John Benjamins. https://doi.org/10.1075/pb.iii.6.03dir

Dirven, René. 2003. In search of conceptual structure. Five milestones in the work of Günter Radden. In Cuyckens, Herbert & Berg, Thomas & Dirven, René & Panther, Klaus-Uwe (eds.), *Motivation in language. Studies in honor of Günter Radden*, xiii–xvi. Amsterdam: John Benjamins.

Dirven, René & Radden, Günter. 2007. *Cognitive English grammar*. Amsterdam: John Benjamins.

Divjak, Dagmar. 2010. *Structuring the lexicon: A clustered model for near-synonymy*. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110220599

Divjak, Dagmar & Fieller, Nick. 2014. Cluster analysis. Finding structure in linguistic data. In Glynn, Dylan & Robinson, Justyna (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 405–441. Amsterdam: John Benjamins.

Dormann, Carsten F. & Elith, Jane & Bacher, Sven & Buchmann, Carsten & Carl, Gudrun & Carré, Gabriel & García Marquéz, Jaime R & Gruber, Bernd & Lafourcade, Bruno & Leitão, Pedro J. & Münkemüller, Tamara & McClean, Collin & Osborne, Patrick E. & Reineking, Björn & Schröder, Boris & Skidmore, Andrew K. & Zurell, Damaris & Lautenbach, Sven. 2012. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 35. 001–020 http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x

Faraway, Julian J. 2016. *Extending the linear model with R generalized linear, mixed effects and nonparametric regression models*, 2nd edition. London: Taylor & Francis.

Geeraerts, Dirk & Grondelaers, Stefan, & Bakema, Peter. 1994. *The structure of lexical variation: Meaning, naming and context*. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110873061

Geeraerts, Dirk & Stefan Grondelaers & Dirk Speelman. 1999. *Convergentie en divergentie in de Nederlandse woordenschat*. Amsterdam: Meertens Instituut.

Gelman, Andreq & Hill, Jennifer. 2007. *Data analysis using regression and multilevel / hierarchical models*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511790942

Glynn, Dylan. 2009. Polysemy, syntax, and variation. A usage-based method for Cognitive Semantics. In Evans, Vyvyan & Pourcel, Stéphanie (eds.), *New directions in cognitive linguistics*, 77–104. Amsterdam: John Benjamins.

Glynn, Dylan. 2010a. Synonymy, lexical fields, and grammatical constructions. A study in usage-based cognitive semantics. In Schmid, Hans-Jörg & Handl, Susanne (eds.), *Cognitive foundations of linguistic usage-patterns: Empirical studies*, 89–118. Berlin: Mouton de Gruyter.

Glynn, Dylan. 2010b. Corpus-driven Cognitive Semantics. Introduction to the field. In Glynn, Dylan & Fischer, Kerstin (eds.), *Quantitative cognitive semantics. Corpus-driven approaches*, 1–42. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110226423.1

Glynn, Dylan. 2010c. Testing the hypothesis. Objectivity and verification in usage-based cognitive semantics. In Glynn, Dylan & Fischer, Kerstin (eds.), *Quantitative cognitive semantics. Corpus-driven approaches*, 239–270. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110226423.239

Glynn, Dylan. 2014a. Polysemy and synonymy. Corpus method and cognitive theory. In Glynn, Dylan & Robinson, Justyna A. (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 7–38. Amsterdam: John Benjamins.

Glynn, Dylan. 2014b. Correspondence analysis. Exploring data and identifying patterns. In Glynn, Dylan & Robinson, Justyna A. (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 443–485. Amsterdam: John Benjamins.

Glynn, Dylan. 2014c. The many uses of *run*. Corpus methods and cognitive sociosemantics. In Glynn, Dylan & Robinson, Justyna A. (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 117–144. Amsterdam: John Benjamins.

Glynn, Dylan. 2016a. Quantifying polysemy. Corpus methodology for prototype theory. *Folia Linguistica* 50. 413–448.

Glynn, Dylan. 2016b. Semasiology and onomasiology. Empirical questions between meaning, naming and context. In Daems, Jocelyne & Zenner, Eline & Heylen, Kris & Speelman, Dirk & Cuyckens, Hubert (eds.), *Change of paradigms. New paradoxes: Recontextualizing language and linguistics*, 47–79. Berlin : Mouton de Gruyter.

Glynn, Dylan & Fischer, Kerstin (eds.). 2010. *Quantitative cognitive semantics. Corpus-driven approaches*. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110226423

Glynn, Dylan & Robinson, Justyna A. (eds.). 2014. *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins.

Gorman, Kyle. 2009. Hierarchical regression modeling for language research. http://repository.upenn.edu/ircs_reports/202

Greenacre, Michael. 2007. *Correspondence analysis in practice*. London: Chapman & Hall.

Gries, Stefan Th. 2003. *Multifactorial analysis in corpus linguistics: A study of particle placement*. London: Continuum.

Gries, Stefan Th. 2014. Frequency tables. Tests, effect sizes, and explorations. In Glynn, Dylan & Robinson, Justyna A. (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 365–389. Amsterdam: John Benjamins.

Gries, Stefan Th. & Divjak, Dagmar. 2009. Behavioral profiles: A corpus-based approach towards cognitive semantic analysis. In Evans, Vyvyan & Pourcel, Stéphanie (eds.), *New directions in cognitive linguistics*, 57–75. Amsterdam: John Benjamins. http://dx.doi.org/10.1075/hcp.24.07gri

Gries, Stefan Th. & Stefanowitsch, Anatol (eds.). 2006. *Corpora in cognitive linguistics. Corpus-based approaches to syntax and lexis.* Berlin: Mouton de Gruyter.

Grondelaers, Stef & Geeraerts, Dirk & Speelman, Dirk. 2007. A case for a cognitive corpus linguistics. In Gonzalez-Marquez, Monica & Mittleberg, Irene & Coulson, Seana & Spivey, Micheal J. (eds.), *Methods in cognitive linguistics*, 149–169. Amsterdam: John Benjamins. http://dx.doi.org/10.1075/hcp.18.12gro

Hanks, Patrick. 1996. Contextual dependency and lexical sets. *International Journal of Corpus Linguistics* 1. 75–98. http://dx.doi.org/10.1075/ijcl.1.1.06han

Heylen, Kris. 2005. A quantitative corpus study of German word order variation. In Kepser, Stephan & Reis, Marga (eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives*, 241–264. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110197549.241

Husson, Francois & Josse, Julie & Le, Sebastiwn & Mazet, Jeremy. 2012. FactoMineR: Multivariate exploratory data analysis and data mining with R. R package version 1.19.

Krawczak, Karolina. 2014. Shame, embarrassment, and guilt. Corpus evidence for the cross-cultural structuring of social emotions. *Poznań Studies in Contemporary Linguistics* 50(4). 441–475.

Krawczak, Karolina. 2015. Epistemic stance predicates in English: A quantitative corpus-driven study of subjectivity. In Glynn, Dylan & Sjölin, Mette (eds.), *Subjectivity and epistemicity. Stance strategies in discourse and narration*, 303–328. Lund: Lund University Press.

Krawczak, Karolina. 2018. Reconstructing social emotions across languages and cultures A multifactorial account of the adjectival profiling of SHAME in English, French, and Polish. *Review of Cognitive Linguistics* 16(2). 454–493.

Krawczak, Karolina & Glynn, Dylan. 2015. Operationalizing mirativity: A usage-based quantitative study of constructional construal in English. *Review of Cognitive Linguistics* 13(2). 253–282. http://dx.doi.org/10.1075/bct.92.04kra

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*. Vol. 1. *Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 1988. A usage-based model. In Rudzka-Ostyn, Brygida (ed.), *Topics in cognitive linguistics*, 127–161. Amsterdam: John Benjamins. http://dx.doi.org/10.1075/cilt.50.06lan

Langacker, Ronald. 1991. *Foundations of Cognitive Grammar.* Vol. 2. *Descriptive applications*. Stanford: Stanford University Press.

Langacker, Ronald W. 1999. *Grammar and conceptualization*. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110800524

Langacker, Ronald W. 2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.

Langacker, Ronald W. 2009. *Investigations in Cognitive Grammar.* Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110214369

Langacker, Ronald W. 2013. *Essentials of Cognitive Grammar*. Oxford: Oxford University Press.

Lindstromberg, Seth. 2010. *English prepositions explained.* Amsterdam: John Benjamins. http://dx.doi.org/10.1075/z.157

Meyer, David & Zeileis, Achim & Hornik, Kurt. 2006. The strucplot framework: Visualizing multi-way contingency tables with vcd. *Journal of Statistical Software* 17. 1–48. http://dx.doi.org/10.18637/jss.v017.i03

Nenadic, Oleg & Greenacre, Michael. 2007. Correspondence analysis in R, with two- and three-dimensional graphics: The ca package. *Journal of Statistical Software* 20. 1–13.

O'Dowd, Elisabeth M. 1998. *Prepositions and particles in English. A discourse-functional account*. Oxford: Oxford University Press.

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.Rproject. org/

Radden, Günter. 1978. Can 'area' be taken out of the waste-basket? In Abraham, Werner (ed.), *Valence, semantic case and grammatical relations*, 327–338. Amsterdam: John Benjamins.

Radden, Günter. 1981. Die übertragenen Bedeutungen der englischen Raumprapositionen. In Radden, Günter & Dirven, René (eds.), *Kasusgrammatik und Fremdsprachendidaktik*, 133–179. Trier: Wissenschaftlicher Verlag.

Rudzka-Ostyn. Brygida. 1989. Prototypes, schemas, and cross-category correspondences: The case of *ask*. *Linguistics* 27. 613–661. http://dx.doi.org/10.1515/ling.1989.27.4.613

Rudzka-Ostyn, Brygida. 2003. *Word power: Phrasal verbs and compounds. A cognitive approach*. Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110197235

Shimodaira, Hidetoshi. 2004. Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling. *Annals of Statistics* 32. 2616–2641. http://dx.doi.org/10.1214/009053604000000823

Speelman, Dirk. 2014. Logistic regression. A confirmatory technique for comparisons in corpus linguistics. In Glynn, Dylan & Robinson, Justyna A. (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 487–503. Amsterdam: John Benjamins.

Suzuki, Ryota & Shimodaira, Hidetoshi. 2011. pvclust: Hierarchical clustering with p-values via multiscale bootstrap resampling. R package version 1.2-2.

Szmrecsanyi, Benedikt. 2010. The English genitive alternation in a cognitive sociolinguistics perspective. In Geeraerts, Dirk & Kristiansen, Gitte & Peirsman, Yves (eds), *Advances in Cognitive Sociolinguistics*, 141–166. Berlin: De Gruyter Mouton.

Taylor, John. 2002. *Cognitive grammar*. Oxford: Oxford University Press. http://dx.doi.org/10.1017/S0272263105230292

Thompson, Laura A. 2009. *S-PLUS (and R) Manual to Accompany Agresti's Categorical Data Analysis (2002).* Available online at www.home.comcast.net/~lthompson221/Splusdiscrete2.pdf

Authors' addresses:

Karolina Krawczak
Faculty of English, Adam Mickiewicz University, Poznań
al. Niepodległości 4
61-874 Poznań, Poland
E-mail: karolina@wa.amu.edu.pl

Dylan Glynn
Université Paris 8 Vincennes-Saint-Denis
2 rue de la Liberté - 93526 Saint-Denis, France
E-mail: dsg.up8@gmail.com

OPERATIONALISIERUNG KONZEPTUELLER PROFILIERUNG. EINE KORPUSBASIERTE STUDIE VON KONSTRUKTIONEN MIT KOGNITIVEN UND KOMMUNIKATIVEN PRÄDIKATEN

Die vorliegende Studie versucht zu veranschaulichen, dass die Erstellung eines sgn. Verhaltensprofils (Behavioural Profile) und insbesondere die multifaktorielle Merkmalsanalyse (Multifactorial Usage-Feature Analysis, Geeraerts et al. 1994; Gries 2003) zur quantitativen Beschreibung von lexiko-grammatischer konzeptueller Profilierung (Langacker 1987, 1999) eingesetzt werden kann. Es wird die Alternanz zwischen of und about als Komplementen bei kognitiven und kommunikativen Prädikaten untersucht. Die Untersuchung umfasst die Prädikate know, speak, talk und think verteilt auf die zwei genannten Präpositionen im britischen und amerikanischen Englisch. Insgesamt wurde ein Korpus von 700 Belegen analysiert, das Annotationsschema beruht auf bereits bestehenden Untersuchungen in diesem Gebiet (Radden 1981; Rudzka-Ostyn 2003; Dirven 2003; Lindstromberg 2010). Durch den Einsatz einer Kombination aus logistischer Regression mit gemischten Effekten, multipler Korrespondenzanalyse and loglinearer Analyse konnte erfolgreich ein Verhaltensprofil für die Alternanz identifiziert werden, das als Operationalisierung der gegensätzlichen Interpretationen betrachtet werden kann. Obwohl sich deutliche Profile zeigen, sind weitere Untersuchungen nötig, um adäquate Mittel für die Trennung des konzeptuellen Beitrags des Prädikats und des Komplements zu finden.

Schlüsselwörter: Prädikat; Präpositionalkomplement; konzeptuelle Profilierung; Korpus; multivariate Analysemethoden.