

Agnieszka Kaleta
University of Lodz
Lodz, Poland

CORPUS-BASED METHODS IN COGNITIVE SEMANTICS:

The case of English clausal complementation

ABSTRACT The paper is concerned with the use of corpus-based methods to deal with issues in cognitive semantics. It presents an overview of the main strands of research combining corpus-based methodology with the theoretical perspective of cognitive linguistics (i.e. behavioral profile approach, collocation analysis). In particular, the paper presents an attempt to assess the applicability of corpus-based methods to studying schematic constructional meanings. Using highlights from Kaleta's 2014 research on English clausal complementation, it pinpoints the strengths and limitations of corpus methodology as a tool for investigating the meanings of abstract syntactic categories. The paper emphasizes the importance of corpora and the methods associated with them for descriptive analysis of syntactic constructions. For example, it shows how the method of distinctive collexeme analysis can be effectively applied to investigate subtle differences in meaning between nearly-synonymous gerundive and infinitival constructions. However, the paper also highlights the limitations of the corpus-based methods when it comes to explanatory analysis. As has been argued, the problematic areas include the high degree of abstractness of syntactic meanings, on the one hand, and the lack of direct correlations between corpus frequencies and abstract mental representations of language, on the other.

1. COGNITIVE SEMANTICS, THE USAGE-BASED MODEL AND CORPUS LINGUISTICS: SOME INTRODUCTORY REMARKS

For almost four decades of its presence on the linguistic scene, cognitive linguistics (henceforth CL) has managed to establish itself as a viable theoretical framework for linguistic analysis. This is evidenced by an ever increasing community of linguists categorizing themselves as ‘cognitive’ and an ever growing number of cognitive linguistic events ranging from small scale, regional conferences to large international congresses gathering both seasoned and novice researchers from all parts of the world. Throughout these four decades of its persistent development, CL (in all its shapes and forms) has brought to our attention many non-trivial aspects of language structure and use, which have not been given due priority in mainstream linguistic research (e.g. in formal approaches).¹ Most notably, cognitive linguistics has put meaning in the center of linguistic enquiry arguing that “meaning is what language is all about” (Langacker 1987: 12). An important implication of this view is that meaning is not an exclusive property of lexical items but that it is also an inherent property of syntactic structures of different degrees of complexity or specificity, including the most abstract syntactic patterns such as the ditransitive or dative construction.

However, the conception of meaning, as entertained by cognitivists is quite distinct from the conception of meaning as held by the advocates of formal approaches or even by functionalists. More specifically, cognitive linguistics emphasizes the conceptual and imaginistic character of meaning. Central to this conception is the notion of construal, understood as our ability to portray the same situation in alternate ways (cf. Langacker 1987: 110). Importantly, rather than being an autonomous faculty in the human mind, linguistic construal is a part of general human cognitive processes, such as those involved in the processing of knowledge, perception or motor activities. Thus, seen from this perspective, meaning is “a mental phenomenon that must eventually be described with reference to cognitive processing” (Langacker 1987: 99).

This conception of language, and linguistic semantics in particular, poses a significant methodological challenge, that is, it raises important questions concerning the methods or procedures with which meanings, as defined by CL, could be effectively studied in a systematic and verifiable way. Much of the research in CL has been informed by individual researchers’ intuitive judgments and by theory-internal considerations. However, there is also a substantial body of experimental research — apart from psycholinguistic experimentation based on elicited data, we can observe an increased interest in corpora and analytical methods associated with them, which hold the promise of ‘scientific objectivism’ and verifiability (cf. Gries 2006). The

Keywords

distinctive collexeme analysis, collostructional analysis, complement constructions, corpus-based methods

¹ The term cognitive linguistics is used here broadly, as a cover term for all strands of cognitive research, including Cognitive Grammar (Langacker 1987, 1991, 2008) as well as constructional approaches (cf. Goldberg 1995, 2006).

interest in corpora can also be attributed to the usage-based thesis of CL, which assumes that mental representations of language arise as a function of language use (rather than being a product of an autonomous mental faculty). In other words, mental grammar is formed by an abstraction or generalization over specific utterances (Barlow & Kemmer 2000; Langacker 2000; Tomasello 2003; Bybee 2006). Langacker (1987: 46) specifies the mental representation of a usage-based grammar as follows:

The grammar lists the full set of particular statements representing a speaker's grasp of linguistic conventions, including those subsumed by general statements. Rather than thinking them an embarrassment, grammarians regard particular statements as the matrix from which general statements (rules) are extracted.

Given this emergentist, non-reductive and bottom-up structure of language, observational data, such as those that can be obtained from linguistic corpora, appear to represent an important source of data for cognitive linguists. In particular, usage-based linguists tend to emphasize the importance of frequency data for the analysis of grammar. There is a wealth of recent results indicating that frequency of occurrence has an enormous impact on the development of linguistic structure and on language use (cf. Bybee 2006, 2010). Bybee (2006: 711) describes the connection between grammar and language use in the following way:

While all linguists are likely to agree that grammar is the cognitive organization of language, a usage-based theorist would make the more specific proposal that grammar is the cognitive organization of one's experience with language. [...] Certain facets of linguistic experience, such as the frequency of use of particular instances of constructions, have an impact on representation that we can see evidenced in various ways, for example, in speakers' recognition of what is conventionalized and what is not, and even more strikingly in the nature of language change. The proposal [...] is that the general cognitive capabilities of the human brain, which allow it to categorize and sort for identity, similarity, and difference, go to work on the language events a person encounters, categorizing and entering in memory these experiences. The result is a cognitive representation that can be called a grammar. This grammar, while it may be abstract, since all cognitive categories are, is strongly tied to the experience that a speaker has had with language.

The main goal of the present paper is to explore the question of how and to what extent the 'non-objectivist' framework of cognitive linguistics can benefit from 'objectivist', frequency-based methods. More specifically, the main focus of the paper is on constructional meanings, i.e. the meanings of abstract syntactic constructions, which appear to pose the greatest meth-

odological challenge for the combined corpus-cognitive approach. The paper is structured as follows. Section 2 presents a brief (and necessarily selective) overview of the state of the art of corpus-cognitive research. Section 3, on the other hand, focuses on one specific aspect of English grammar — English clausal complementation — making it a test case for assessing the extent to which corpus-based methods provide a useful methodological framework for research into constructional meanings. It begins with a brief outline of the previous work on English complement constructions (section 3.1.), which is followed by a section presenting the results of Kaleta's (2014) corpus-based study on quasi-synonymous constructions with verbs accepting two different complement types (e.g. *start_to* vs. *start_ing*). These results are subjected to critical examination in Section 4, which discusses the merits and limitations of the frequency-based approach to the meanings of abstract syntactic constructions.

2. COGNITIVE LINGUISTICS MEETS CORPUS LINGUISTICS: AN OVERVIEW OF CORPUS-BASED APPROACHES TO ISSUES IN COGNITIVE SEMANTICS

Corpora usage has been present in cognitively oriented research from its very inception. The early studies have been concerned mainly with lexical phenomena, exploring the ways in which various lexical concepts or sets of semantically closely related concepts carve up (structure) conceptual space (cf. Dirven et al. 1982; Dirven & Taylor 1988; Geeraerts 1993 a, b; Rudzka-Ostyn 1995, 1996). These studies relied mainly on simple counts, raw frequencies or percentages. Over time, however, more and more attention has been given to the issue of statistical validation or assessment of raw frequencies. Consequently, the methods used by usage-based linguists have become more advanced and sophisticated in terms of statistical analysis. One of the methods allowing for incorporation of different statistical techniques is the so-called 'behavioral profile approach', developed mainly to study issues in lexical semantics (Divjak 2006; Divjak & Gries 2006; Gries & Divjak 2009). It begins with the retrieval of the search items from the corpus in the form of concordance lines, which are annotated for many different features: morphological, syntactic, semantic, and sometimes also discourse-pragmatic. Next, all these features are together subjected to statistical evaluation by various statistical techniques such as pairwise differences of percentages, some correlational statistics or the most recommended for this types of research cluster analysis. All these statistical procedures are exploratory in nature, that is they allow to detect structure in complex, multidimensional datasets. For example, Divjak and Gries's (2006) study of nine Russian verbs meaning 'to try' employs hierarchical cluster analysis to determine the number of semantic

classes that the synonyms fall into. Gries and Divjak (2009), on the other hand, apply behavioral profile approach to a comparative study of English and Russian near-synonyms, showing how the statistical technique of computing differences between percentages can provide information on subtle cross-linguistic formal/semantic differences that underlie lexical choices in different languages. Examples of behavioral profile studies of polysemy include Gries's (2006) analysis of the many senses of *to run* and Berez and Gries's (2009) analysis of the highly polysemous *to get*.

A similar approach can be found in studies dealing with grammatical constructions, and, in particular, with syntactic alternations. It should be emphasized that most corpus-based cognitive work on syntax has worked from the idea of a construction, as used in both Construction Grammar (Goldberg 1995) and Cognitive Grammar (Langacker 1991, 2002). Seen from this perspective a construction is any linguistic expression that is directly associated with a particular meaning. As argued by Golberg (2006: 5),

Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency.

Language is thus viewed as a vast inventory of successively more abstract constructions (form-meaning pairings), ranging from words to various multi-word expressions (variable or fully fixed) to abstract syntactic templates such as the passive or the ditransitive construction.

This idea of the construction underlies most corpus-cognitive work on syntax, one strand of which resembles behavioral profile approach, as briefly outlined above. More specifically, it involves coding instances of quasi-synonymous constructions for a variety of semantic and formal features which are subsequently submitted to statistical evaluation with a view to determining the factors responsible for the choice of one construction over the other one. This type of analysis can be found in Gries's (2001) study of alternations involving shifts in particle placement (e.g. *She put the book down vs. She put down the book*). Using a regression like method called linear discriminant analysis, Gries attempts to determine the factors that best predict the speaker's constructional choices. Other examples of this line of research include Grondelaers, Speelman, and Geeraerts (2002), Bresnan et al. (2007), Szmrecsanyi (2010), and Levshina (2012).

Another method dedicated specifically to studying syntactic constructions is collostructional analysis, as developed by Stefanowitsch and Gries (2003), Gries and Stefanowitsch (2004). It is a collocation-based method, which is aimed at studying semantic properties of schematic (abstract) con-

structions by examining how the words filling constructional slots fall into semantic classes. The main idea behind this approach is that:

[...] a word may occur in a construction if it is semantically compatible with the meaning of the construction (or, more precisely, with the meaning assigned by the construction to the particular slot in which the word appears). (Stefanowitsch & Gries 2003: 213).

The method measures the degree of attractions that words (the so-called collexemes) exhibit to constructional slots in which they occur. The most common statistics used for this type of analysis is Fisher-Yates exact test. Collostructional analysis subsumes three different, albeit closely related methods:

- collexeme analysis which deals with measuring the degree of association between a particular lemma (a collexeme) and a slot in a construction (cf. Stefanowitsch & Gries 2003);
- distinctive collexeme analysis which is specifically geared towards investigating semantic properties of pairs of semantically similar grammatical constructions (syntactic alternations). Its main aim is to identify lexemes that exhibit a significantly higher preference for one member of the pair than for the other (cf. Gries & Stefanowitsch 2004);
- covarying collexeme analysis which serves to investigate the interrelations between the words filling two different slots in the same construction (cf. Stefanowitsch & Gries 2005).

To date, collostructional analysis has been employed to investigate the semantic range of various syntactic constructions, and in particular, to spell out subtle semantic differences between syntactic quasi-synonyms. Examples of this type of analysis include Wulf (2006), Hampe (2011), Hoffman (2011), and Kaleta (2014). Noteworthy is also Guilquin's (2010) extension of distinctive collexeme analysis to multiple distinctive collexeme analysis which distinguishes between several closely related constructions rather than just two alternative forms. Collostructional analysis has also been used to study diachronic development of nearly synonymous constructions (cf. Hilpert 2006, 2008; Coleman 2009; Noël & Coleman 2010).

Succinct and selective as this overview may be, it shows that quantitative methods have been enjoying considerable popularity among cognitively-oriented researchers, who do not give up on their attempts to introduce new analytical techniques and refine the old ones. The two methods, as outlined above (i.e. collostructional analysis and behavioral profile approach) are illustrated in Section 3 with examples from Kaleta's (2014) corpus-cognitive study of quasi-synonymous complement constructions with verbs that accept two (or more) different complement types. However, first a brief look is taken at how complement constructions have been described and analyzed in cognitive linguistic literature.

3. ENGLISH CLAUSAL COMPLEMENTATION

This section begins with a brief outline of the main theories of English post-verbal complement constructions: the infinitive, the gerund, and finite *that* clauses. A special emphasis is given to the functional/cognitive theories according to which the distribution of complement categories is not random or arbitrary, but motivated by schematic meanings of these categories and (or) by the iconic principle of conceptual distance between the main clause event and the complement event.

3.1. COGNITIVE APPROACHES TO ENGLISH CLAUSAL COMPLEMENTATION

Much cognitive research into the semantics of complement constructions has been influenced by functionalist approaches, especially by work of Bolinger (1968, 1974), Givón (1980, 1985), and Wierzbicka (1988). For example, Langacker (1991: 445–446) agrees with Wierzbicka that the *to*-infinitive expresses futurity, while the gerund codes temporal overlap between the main clause event and the complement clauses event, with the reservation that these may be only prototypical characteristics of these two complement types. However, in line with his theory of Cognitive Grammar, Langacker goes on to propose a more schematic characterization of these two complement types. He argues that the gerund has the same function as the progressive construction, i.e. it imposes a restricted, internal perspective on the complement process. The *to*-infinitive, on the other hand, encodes the PATH image-schema, that is, it construes the complement scene as a path leading to its completion (Langacker 1991: 446). The PATH theory, of course, has historical grounding — it draws on the view that the infinitival *to* derives from the prepositional *to* meaning ‘toward’. While the idea that the infinitival complementation is motivated (in one way or another) by the PATH imagery has been widely accepted by researchers working within the cognitive linguistics framework, the simultaneity ascribed to the gerundive complement has been more contentious. Another perspective is that the function of the gerundive complement is merely the general process of conceptual reification, which involves construing the complement scene as an abstract ‘thing’ (Duffley 2006; Kaleta 2014). Thus, from this perspective, the gerundive *-ing* is much like the nominalizing *-ing* (a view that also has its etymological roots).

As for finite *that* clauses, the dominant view is that they represent episodic judgments and conceptions of reality rather than actual occurrences (cf. Langacker 2008). Another (complementary) view appeals to the notion of iconicity, and specifically, to the idea that the choice of the complement type by a given verb depends on the degree of conceptual distance between the main clause event and the complement event. This distance correlates with

the phonological distance between the two clauses — it is greatest with *that* and shortest with bare infinitive complements (cf. Givón 1980; Verspoor 2000). Thus, the verbs that refer to pure reasoning typically select *that* clauses, whereas the verbs that suggest some direct interaction between the main clause subject and the complement event take the infinitival complement (e.g. *make, see*).

The theories of English complementation, as sketched above, have been derived largely through introspective analyses seeking to determine the connections between the shared semantics of different types of verbs and the unitary semantic value of their respective complementation patterns. While it has yielded a number of interesting observations, this approach seems to lose its appeal when it comes to more subtle semantic differences like those that accompany alternations where the same matrix verb accepts two or more complement types, with no obvious differences in meaning. A corpus-based approach appears to represent a promising alternative to introspection-based analyses in such cases. This point is further developed in the next section, which looks specifically at the applications of corpus-based methods to syntactic alternations between *-ing* and *to*-infinitive complements.

3. 2. A CORPUS-BASED APPROACH TO ENGLISH COMPLEMENT CONSTRUCTIONS AT WORK

This section illustrates the application of corpus based-methods to issues in cognitive semantics using examples from Kaleta's (2014) research on nearly synonymous complement constructions. First, a look is taken at the method of distinctive collexeme analysis, which has been applied to non-finite complementation of English verbs of aspects and verbs of like (and dislike). Next, the results of the behavioral profile approach are illustrated with some examples from the studies on verbs accepting both non-finite complements and finite *that* clauses.

3.2.1. DISTINCTIVE COLLEXEME ANALYSIS OF TO- INFINITIVE AND -ING ALTERNATIONS²

Kaleta (2014) has applied the method of distinctive collexeme analysis to investigate subtle semantic differences between infinitival and gerundive constructions, as they occur with aspectual verbs (*begin, start, cease, continue*) and verbs of like and dislike (*like, love, hate, prefer, dread*). The analysis has revealed some distinctive distributional preferences of the nearly-synonymous constructions. By way of illustration, let us consider the 20 top collexemes of *start to Verb* and *start Verb ing*. The collexemes are listed in descending order of their strength of attraction to the infinitival and gerundive slots (the smaller the Fisher's p value the higher the strength of attraction of a given verb to the construction).

² All the analyses presented in this section have been carried out using the British National Corpus and Sketch Engine software.

Figure 1. Twenty most distinctive collexemes of *start to Verb* and *start Verb ing*

START TO	Fisher exact	START_ING	Fisher exact
<i>feel</i>	6,54E-022	<i>talking</i>	5,00E-023
<i>emerge</i>	5,66E-011	<i>shipping</i>	3,08E-008
<i>appear</i>	4,52E-010	<i>doing</i>	3,24E-008
<i>become</i>	3,16E-009	<i>training</i>	4,58E-007
<i>rise</i>	1,62E-008	<i>thinking</i>	1,55E-006
<i>fall</i>	8,96E-007	<i>shooting</i>	1,66E-006
<i>climb</i>	1,43E-006	<i>trying</i>	6,19E-006
<i>rain</i>	1,54E-006	<i>calling</i>	1,03E-005
<i>grow</i>	5,01E-006	<i>playing</i>	2,00E-005
<i>show</i>	6,34E-006	<i>paying</i>	1,21E-004
<i>change</i>	7,12E-006	<i>shouting</i>	1,32E-004
<i>turn</i>	8,35E-006	<i>saving</i>	1,45E-004
<i>get</i>	9,72E-006	<i>dancing</i>	2,57E-004
<i>return</i>	1,17E-005	<i>selling</i>	2,93E-004
<i>pick</i>	1,78E-005	<i>manufacturing</i>	7,40E-004
<i>fade</i>	2,59E-005	<i>yelling</i>	7,40E-004
<i>move</i>	3,24E-005	<i>planning</i>	1,00E-003
<i>relax</i>	5,72E-005	<i>asking</i>	1,28E-003
<i>crumble</i>	5,72E-005	<i>painting</i>	1,39E-003
<i>shake</i>	9,83E-005	<i>throwing</i>	1,65E-003

As can be seen, *start to Verb* and *start Verb ing* show clearly different collocational preferences: *start to Verb* preferentially co-occurs with state or change-of state verbs, while *start Verb ing* significantly attracts action verbs. The only exception is the collexeme *thinking*. However, it occurs in the gerundive slot only in the agentive sense ‘to consider sth in one’s mind.’ Similar results have been obtained for other aspectual verbs, which may be taken as indicative of a stable semantic contrast between the two complement types. Given this distribution, Kaleta (2014) concludes that the infinitival constructions (*start to*, *begin to*, *cease to*) prototypically denote a change or transition into a new state, while their gerundive counterparts profile the onset, continuation or cessation of ongoing processes.

The question that arises at this point is how this distribution fits into the broader picture of *to*-infinitive and *ing* complements, as outlined in the section above, or, in other words, how it can be accounted for in a cognitively plausible way. It is clear from the distinctive collexeme analysis that the temporal notions of futurity and simultaneity, as postulated by Wierzbicka (1988), are insufficient to deal with the alternations in question. Thus, a more general characterization is needed and Langacker’s (1991, 2008) description of the *to*-infinitive in terms of the PATH image schema appears to fulfill this requirement. As argued in Kaleta

(2014), the notion of a change and that of motion towards a goal, as encoded in the PATH schema, are linked metaphorically, that is via the (well documented) MOTION IS CHANGE conceptual metaphor. In other words, the ‘change’ meaning constitutes a metaphorical extension or elaboration of the schematic notion of the PATH, which provides a unitary, schematic characterization of the infinitival complement. Of course the notion of futurity does play an important role in the distribution of the infinitival complement, and it is clearly compatible with the PATH schema in that moving towards a target destination is correlated with ‘a shift’ in time (i.e. a target location is ‘later’ than a source location). However, the case of aspectual constructions shows that ‘futurity’ is not the only and definitely not a self sufficient explanation for the distribution of the infinitival complement and that this explanation must be sought at a deeper level of cognitive schemas and metaphorical processes. The same conclusion seems to extend to the gerundive constructions. As has been observed, these constructions tend to evoke the notion of an ongoing process. However, this notion is not an inherent property of the gerundive construction, as has been postulated by some linguists, but rather a result of the process of re-temporalization of the complement event. That is, aspectual verbs appear to have a function similar to that of the *be* operator in the progressive construction — they re-temporalize the atemporal (nominal) construal imposed by the gerundive complement, which gives rise to the notion of ‘ongoingness’ or a ‘progressing’ event.

As another example, let us consider the results of a distinctive collexeme analysis of *like to Verb* and *like Verb ing*. The tables below present the fifteen most distinctive collexemes of the two constructions:

LIKE_TO	Fisher exact	LIKE_ING	Fisher exact
<i>think</i> (293:11)	1,1E-094	<i>being</i> (311:759)	1,27E-28
<i>know</i> (126:11)	1,2E-034	<i>living</i> (14:41)	1,53E-13
<i>see</i> (191:51)	2,1E-031	<i>doing</i> (114:105)	6,10E-12
<i>call</i> (77:6)	3,4E-022	<i>going</i> (92:73)	6,72E-07
<i>keep</i> (84:14)	9,3E-019	<i>talking</i> (52:43)	7,00E-05
<i>hear</i> (59:11)	7,1E-013	<i>watching</i> (33:30)	2,91E-04
<i>feel</i> (45:7)	6,4E-011	<i>working</i> (48:38)	3,34E-04
<i>believe</i> (18:0)	2,2E-007	<i>paying</i> (4:10)	6,11E-04
<i>admit</i> (23:3)	1,7E-006	<i>meeting</i> (5:10)	1,39E-03
<i>show</i> (23:4)	6,6E-006	<i>playing</i> (46:33)	2,55E-03
<i>consider</i> (20:3)	1,5E-005	<i>driving</i> (5:9)	3,55E-03
<i>pretend</i> (11:0)	8,5E-005	<i>wearing</i> (18:17)	4,60E-03
<i>remind</i> (11:0)	8,5E-005	<i>shopping</i> (2:6)	6,08E-03
<i>spend</i> (22:7)	2,9E-004	<i>dealing</i> (7:9)	1,19E-02
<i>describe</i> (8:0)	1,1E-003	<i>losing</i> (3:6)	1,41E-02

Figure 2. Fifteen most distinctive collexemes of *like to Verb* and *like Verb ing*

As this analysis indicates, *like to Verb* preferentially co-occurs with cognition verbs (*think, know, believe, consider*), sense perception verbs (*see, hear, feel*) and speech act verbs (*admit, describe, maintain*). The gerundive construction, on the other hand, most strongly attracts state verbs *being* and *living*, with all the other slots being filled with common action verbs.

Interesting as these results may be from the descriptive perspective, they do not tell us much about the semantic factors motivating the choice of the gerund or the *to*-infinitive. However, a closer look into the contextualized uses of the significant collexemes of the two constructions does reveal some tendencies that confirm the distinction made in the previous literature between the infinitival and gerundive complements of emotion verbs in terms of the ‘desire’ and ‘enjoyment’ meaning, respectively (cf. De Smet & Cuyckens 2005). Let us consider some examples of the infinitival constructions filled with cognition verbs³:

- (1) She *likes to think* of herself as coming from a higher social class.
- (2) Frenchmen still *like to believe* they're the world's greatest lovers.

As these examples show, *like to Verb* is not confined solely to the enjoyment meaning but introduces an additional, volitional component, which, in the cases under consideration, takes the form of ‘wishful thinking’. Apart from state verbs, the infinitival construction co-occurs with action verbs, in which case the overall meaning of *like to Verb* seems to appeal to the general notion of habituality, i.e. a regular occurrence of an event. Consider the following examples:

- (3) At times I *like to speak* English, but not for a very long time.
- (4) How do you *like to spend* your spare time?
- (5) I *like to show* people I've got guts, although I've been scared at times.

It should be pointed out that the enjoyment meaning, as encoded by the matrix predicates, is still present in the infinitival constructions, i.e. the notions of volitionality and habituality tend to overlap with that of enjoyment, forming a kind of semantic blend. However, when negated, agentive uses of *like to Verb* no longer evoke habituality — they typically refer to a single act of ‘not wanting to do sth’. For example:

- (6) Colonel, I *do not like* to remind you of this, but I saved your son's life.
- (7) Juliet *didn't like* to admit that the same thought crossed her own mind.

³ All the examples included in this section are taken from the British National Corpus.

Gerundive uses contrast with the infinitival ones in that they denote pure ‘enjoyment’ without introducing any additional semantic components. For example:

- (8) She *likes* talking about her brother. She hasn’t many other pleasures.
- (9) [...] I *like* being in the company of a pretty girl, even if she looks only twelve.
- (10) As for me, I *liked* going to tea with Mr. and Mrs. Wilson. I loved the orderliness of their little bungalow, and their well-kept garden.

In sum, while the overall meaning of *like Verb ing* is consistent with the lexical semantics of the matrix predicate, *like to Verb* shows a tendency to evoke additional semantic components, i.e. those of volitionality and habituality. Importantly, these meanings are not random or arbitrary — they appear to be metonymically related to the basic enjoyment meaning of *like*. The connection is straightforward: if we enjoy doing something, we want to do this, and we want to do this often/repeatedly (cf. De Smet & Cuyckens 2005).

While the corpus data provide an insight into the usage and meanings of specific constructions with verbs of like (and dislike), the explanation of this usage requires some analytical reasoning. Specifically, the volitional component, as traced in the infinitival constructions, appears to be linked to the PATH image schema via the common conceptual metaphor GOALS ARE DESTINATIONS. It should be noted that this use of the infinitival complement is not an idiosyncratic fact but rather a confirmation of the general tendency of the *to*-infinitive to combine with verbs which have in their semantic structure the notion of ‘wanting’ (cf. Wierzbicka 1988). As regards the gerundive constructions, the observed tendency of the constructional meanings to coincide with the lexical meanings of the matrix predicates (enjoyment) appears to confirm the hypothesis that the gerundive complement does not contribute any specific semantic content apart from the general function of imposing atemporal, nominal construal on the complement scene.

So far we have seen how collostructional analysis can be applied to the analysis of constructional meanings. However, given that the scope of this method is limited to studying constructions with fixed syntactic slots, more variable complement constructions, like the ones introduced with *that* clauses, have been studied using the method of behavioral profiles, which is illustrated in the following section.

3.2.2. FINITE VS. NON FINITE COMPLEMENTS: A BEHAVIORAL PROFILE APPROACH

The procedure, as followed in Kaleta (2014), consists of two steps: annotating sample concordances of the constructions being studied for a variety of features (formal and semantic), and statistical evaluation of the corpus frequencies of these features with a chi-squared test in order to determine the features that distinctively differentiate one construction from the other one. To exemplify, let us briefly consider two quasi-synonymous uses of *promise*, i.e. *promise* followed by the *to*-infinitive and *promise* followed by the *that* clause with a subject co-referential with the main clause subject (e.g. *He promised to do this* vs. *He promised that he will do this*). Seen from the cognitive linguistic perspective, these two constructions represent a shift from subjective to objective construal in the sense of Langacker (2008: 77–78). The main aim of the study has been to determine the factors responsible for the choice of one perspective over the other one. This analysis has shown that *promise that* is distinctively associated with the presence of a direct object (*promise sb. that*), the first person subject (*I*), and the present simple tense (*I promise*). Unfortunately, the import of this distribution becomes more apparent only in the course of a qualitative evaluation of these significant patterns. This evaluation has revealed that *promise that* is distinguishable from *promise to* mainly on the basis of its emphatic properties. That is, *promise that* tends to impose an emotional and involved perspective on the complement scene, which contrasts with a much more neutral perspective of *promise to*. Consider the following examples:

- (11) Please, Alice. Please *promise* me you'll phone.
- (12) I warn you that I will take steps to prevent you or I *promise* that I will be there on time.

The presence of a strong emotive component in the construal of the complement scene makes the infinitival complement much less likely:

- (13) ?? Please, Alice. Please *promise* me to phone.
- (14) ?? I warn you that I will take steps to prevent you or I *promise* to be there on time.

Finally, let us briefly consider the case of *that* and *ing* alternations, as represented by *admit Verb ing* and co-referential *admit that*. As for the former, the corpus analysis has revealed its significant association with action verbs denoting acts of unlawful or criminal behavior. The construction is also significantly skewed toward the third person subjects (typically personal names)

and the complement event is invariably anterior with respect to the time of the main clause event (admitting sth). The construction is most commonly found in reports of court proceedings. Here are some typical examples:

- (15) In court 23 year old Colin Longmur *admitted* causing a breach of the peace at the couples.
- (16) At Belfast Crown Court, Corry *admitted* murdering Mr Magee [...].
- (17) Michael Sams, 51, has already *admitted* kidnapping 25-year-old Stephanie Slater.

Given this distribution, the meaning of *admit Verbing* can be glossed as 'admit one's blame'. A comparison with a sample concordance of distinctive uses of *admit that* reveals a subtle shift of perspective. That is, the finite construction appears to evoke 'the other oriented' perspective in the sense of Wierzbicka (1988). Specifically, under this construal the complement situation has an 'objective' existence, i.e. it is accessible to other conceptualizers and can be assessed for validity. The subject referent takes the position of someone who accepts (or does not accept) what others think or believe to be true. Thus, the meaning of this construction can be glossed as 'to accept the truth of'. This conclusion has been derived from a qualitative assessment of event types distinctively coded by *admit that*, which include states and agentive events other than acts of unlawful behavior. Compare examples (15)–(17) with the following ones:

- (18) He has put on 4,5 stone, and at 13 stone *admits* he is now overweight.
- (19) And what do you mean, I've met her? I *admit* I came close, the night I delivered flowers.
- (20) For one the Halifax was following others: it *admits* that it paid too much at the wrong time.

Having exemplified how corpus-based methods have been employed to study English complement constructions, in the next section I discuss the capabilities and limitations of the corpus approach to constructional meanings.

4. DISCUSSION AND CONCLUSIONS

A systematic application of corpus-based methods to study constructional meanings has a number of advantages. First and foremost, it provides a detailed insight into the full range of distributional properties of syntactic constructions, allowing an insight into the subtleties of linguistic usage that could

remain inaccessible to purely intuitive analysis. This is of course of particular importance to studying quasi-synonymous constructions, which represent subtle (yet important from both descriptive and theoretical perspective) shifts in meaning. For example, none of the introspection-based approaches have accurately captured the range of use of alternating aspectual constructions (e.g. *begin to Verb* vs. *begin Verb ing*), which, in turn, has led to some misconceptions concerning the semantic function of the gerundive and infinitival complements. It should be clear that collostructional approach has opened new perspectives in the research on nearly-synonymous constructions, enabling a more accurate delineation of their distribution and thereby providing a clearer understanding of their semantic (or conceptual) content.

Another important advantage of corpus-frequencies is their utility in isolating prototypical uses from more peripheral ones. As argued by usage-based linguists prototypicality is closely correlated with frequency of occurrence (cf. Barlow & Kemmer 2000; Schmid 2000). Thus, frequencies of occurrence provide an important clue to distinguishing more typical functions of a construction from the more peripheral ones. In particular, they are an important tool in capturing the gradient or scalar nature of linguistic categories. As pointed out above, this is of particular importance to studies on nearly-synonymous constructions, which, more often than not, are a matter of degree rather than all-or-nothing distinctions. As a matter of fact, all the constructional pairs discussed in the section above are only partially exclusive, that is, they all have subsets of overlapping uses. For example, both *start Verb ing* and *start to Verb* can be used to express agentive events. However, the corpus frequencies clearly show that the former is more likely to denote change-of-state situations, while the latter is typically used to refer to dynamic, ongoing situations.

As vital as it is, descriptive analysis of linguistic usage is not an end in itself. Apart from the important question of 'what language is like', equally important or even more important is the question of 'why language is the way it is'. Thus, the ultimate goal of linguistic research is that of explaining language structure and use. It is this explanatory aspect of linguistic research that poses the greatest challenge to corpus-cognitive approach. The difficulty arises, naturally, from the very conception of language, as proposed by cognitive linguists. It should be clear that a frequency-based approach must encounter difficulties when faced with a conception of language which assumes that grammar is conceptualization and that it involves active construal rather than a passive reflection of an objective world. Furthermore, as cognitively-oriented research shows, the ultimate explanation for language structure and use must be sought at the level of general cognitive structures and operations. This appears to be particularly true of syntactic constructions, whose meanings are characterized by a high degree of schematicity or abstractness. Thus, the essential question that needs to be addressed at this point is whether corpus frequencies do actually reflect

mental representations of language. The answer to this question seems to be in the positive, however with the reservation that the link between frequencies of occurrence and mental representations of linguistic structures is often indirect or at least much less direct than many corpus-based cognitive linguists believe it to be or would like it to be. A distinction has to be made at this juncture between lower-level (more specific) constructions and higher-level, that is, more schematic, constructions. For example, *start Verb to* and *start Verb ing*, as discussed in the section above, represent the more specific end of the constructional continuum (due to the presence of lexical fillers), with the infinitival and gerundive complements representing a higher-order schemata. In the network model of linguistic structure, as proposed by Langacker (2002: 271), the more specific uses are seen as instantiations or elaborations of more schematic constructions, which entails that the former cannot be satisfactorily accounted for without gaining an insight into the latter. The problem that this view creates for a corpus-based approach is that of operationalization of abstract syntactic meanings. This can be clearly seen in the case of the infinitival and gerundive constructions — while there are strong arguments, particularly historical/etymological ones, in favor of the thesis that the English *to*-infinitive is motivated by the PATH imagery and that the function of the gerund is merely that of imposing an atemporal/nominal construal on a scene, it should be clear that structures of this kind do not lend themselves easily to frequency-based verification. The presence of metaphorical and metonymic processes in the structure of these two complement types additionally complicates such verification.

Also, as we have seen, the characterization of finite *that* clauses in terms of epistemic construal and conceptions of reality poses a challenge for a frequency-based approach. This theory has a strong basis in the general distribution of *that* complements, which combine mainly with predicates expressing knowledge or information (cf. Langacker 2008; Wierzbicka 1988). However, the verification of this theory in the context of minimal pair constructions with a single verb alternating between the *that* complement and a non-finite complement is by no means a straightforward task. Of course, corpus data provide a number of important clues to the semantic structures motivating particular constructions, yet these are only 'clues' rather than direct indicators of semantic content. For example, the observed strong tendency of co-referential *promise that* to express emphatic, emotional construal can be taken as indicative of the presence of epistemic component in that the committed (emotional) attitude represents a kind of epistemic attitude or stance. A similar indirectness can be observed in the case of *admit Verb ing* and co-referential *admit that*, as briefly discussed in the section above. While the corpus frequencies provide a reliable picture of the distributional profiles of the two constructions, the explanation of these profiles/patterns requires a more analytical, top-down approach. More specifically, the patterns in question lend themselves to the explanation along

the lines proposed by Langacker (2008: 441–444), who accounts for the general contrast between non-finite and finite complements in terms of ‘occurrences’ and conceptions of reality, respectively. That is, *admit Verb ing*, with its meaning ‘to admit one’s guilt’, profiles the ‘effective’ level (‘occurrences’) — it merely suggests that the subject referent committed a crime or was responsible, in one way or another, for the complement event. *Admit that*, on the other hand, with its propositional meaning ‘to accept the truth of something’ enters into the epistemic domain of conceptions of reality.

The picture that emerges from the foregoing discussion is that syntactic categories (and complement constructions in particular) represent a complex reality, the full understanding of which requires different sources of data and the application of different methods of analysis (both bottom-up and top-down ones). As has been seen, corpus frequencies provide a useful tool for descriptive purposes, yet their explanatory power is subject to certain limitations. The issue of operationalizing abstract/schematic meanings and, in particular, the indirectness of the link between corpus-frequencies and mental representations of language appear to be the biggest challenge for a combined corpus-cognitive approach. Also, the fuzziness and fluidity of linguistic/conceptual categories provide an obstacle for this approach. However, these limitations do not undermine the applicability of frequency based methods to studying constructional meanings. They merely show that a frequency-based approach to semantics may not follow exactly the same logic as empirical work in ‘hard sciences’, that is, it may not achieve the same degree of scientific rigor and objectivity. This, however, is not the fault of the method (which is functional, as far as it goes), but rather a natural consequence of the complex and multifaceted nature of linguistic meanings, which are not subject to the same ‘laws of predictability’ as other aspects of human behavior. This, of course, is not to say that no new improvements or refinements can and should be expected in the methodological field. Quite to the contrary, it seems that more conscious effort should be directed towards methodological progress in the empirical study of linguistic meanings, and, in particular, in corpus-based studies of constructional meanings, as defined by CL.

REFERENCES

- Barlow, Michael, Suzanne Kemmer (eds.) 2000: *Usage-Based Models of Language*. Stanford: CSLI Publications.
- Berez, Andrea L., Stefan Th. Gries 2009: In defense of corpus-based methods: A behavioral profile analysis of polysemous *get* in English. In: Steven Moran, Darren S. Tanner, Michael Scanlon (eds.) 2009: *Proceedings of the 24th Northwest Linguistics Conference* [University of Washington Working Papers in Linguistics, vol. 27]. Seattle: Department of Linguistics, 157–166.

- Bolinger, Dwight 1968: Entailment and the meaning of structures. *Glossa* 2, 119–127.
- Bolinger, Dwight 1974: Concept and percept: Two infinitive constructions and their vicissitudes. In: Nihon Onsei Gakkai (ed.) 1974: *World Papers in Phonetics: Festschrift for Dr. Onishi's Kiju*. Tokyo: Phonetic Society of Japan, 65–91.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina, R. Harald Baayen 2007: Predicting the dative alternation. In: Gerlof Bouma, Irene Krämer, Joost Zwarts (eds.) 2007: *Cognitive Foundations of Interpretation*. Amsterdam: Royal Netherlands Academy of Science, 69–94.
- Bybee, Joan 2006: *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.
- Bybee, Joan L. 2010: *Language, Usage and Cognition*. Cambridge: Cambridge University Press.
- Colleman, Timothy 2009: The semantic range of the Dutch double object construction: A collostructional perspective. *Constructions and Frames* 1, 190–221.
- De Smet, Hendrik, Hubert Cuyckens 2005: Pragmatic strengthening and the meaning of complement constructions: The case of *like* and *love* with the *to*-infinitive. *International Journal of Corpus Linguistics* 33, 3–34.
- Dirven, René, Louis Goossens, Yvan Putseys, Emma Vorlat 1982: *The Scene of Linguistic Action and Its Perspectivization by SPEAK, TALK, SAY and TELL*. Amsterdam, Philadelphia: John Benjamins.
- Dirven, René, John R. Taylor 1988: The conceptualization of vertical space in English: The case of *tall*. In Brygida Rudzka-Ostyn (ed.) 1988: *Topics in Cognitive Linguistics*. Amsterdam, Philadelphia: John Benjamins, 379–402.
- Divjak, Dagmar 2006: Ways of intending: Delineating and structuring near synonyms. In: Stefan Thomas Gries, Anatol Stefanowitsch (eds.) 2006, 19–56.
- Divjak, Dagmar, Stefan Thomas Gries 2006: Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2, 23–60.
- Duffley, Patrick J. 2006: *The English Gerund-Participle: A Comparison with the Infinitive*. Frankfurt am Main: Peter Lang.
- Geeraerts, Dirk 1993a: Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics* 4, 223–272.
- Geeraerts, Dirk 1993b: Generalised onomasiological salience. In: Jan Nuyts, Eric Pederson (eds.) 1993: *Perspectives on Language and Conceptualization* (Special ed. *Belgian Journal of Linguistics* 8). Brussels: Presses Universitaires de Bruxelles, 43–56.
- Gilquin, Gaëtanelle 2010: *Corpus, Cognition and Causative Constructions*. Amsterdam, Philadelphia: John Benjamins.
- Givón, Talmy 1980: The binding hierarchy and the typology of complements. *Studies in Language* 4.3, 333–377.

- Givón, Talmy 1985: Iconicity, isomorphism and non-arbitrary coding in syntax. In: John Haiman (ed.) 1985: *Iconicity in Syntax*. Amsterdam, Philadelphia: John Benjamins, 187–219.
- Goldberg, Adele E. 1995: *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006: *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Gries, Stefan Thomas 2001: A multifactorial analysis of syntactic variation: Particle movement revisited. *Journal of Quantitative Linguistics* 8, 33–50.
- Gries, Stefan Thomas 2006: Corpus-based methods and cognitive semantics: The many senses of *to run*. In: Stefan Thomas Gries, Anatol Stefanowitsch (eds.) 2006, 57–99.
- Gries, Stefan Thomas, Dagmar S. Divjak 2009: Behavioral profiles: A corpus-based approach towards cognitive semantic analysis. In: Vyvyan Evans, Stephanie S. Pourcel (eds.) 2009: *New Directions in Cognitive Linguistics*. Amsterdam, Philadelphia: John Benjamins, 57–75.
- Gries, Stefan Thomas, Anatol Stefanowitsch 2004: Extending collocation analysis: A corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics* 9.1, 97–129.
- Gries, Stefan Thomas, Anatol Stefanowitsch (eds.) 2006: *Corpora in Cognitive Linguistics: Corpus Based Approaches to Syntax and Lexis*. Berlin, New York: Mouton de Gruyter.
- Grondelaers, Stefan, Dirk Speelman, Dirk Geeraerts 2002: Regressing on *er*: Statistical analysis of texts and language variation. In: Anne Morin, Pascal Sébillot (eds.) 2002: *6th International Conference of the Statistical Analysis of Textual Data*. Rennes: Institut National de Recherche en Informatique et en Automatique, 335–346.
- Hampe, Beate 2011: Discovering constructions by means of collocation analysis: The English Denominative Construction. *Cognitive Linguistics* 22.2, 211–245.
- Hilpert, Martin 2006: Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2.2, 243–257.
- Hilpert, Martin 2008: *Germanic Future Constructions: A Usage-Based Approach to Language Change*. Amsterdam, Philadelphia: John Benjamins.
- Hoffmann, Thomas 2011: *Preposition Placement in English: A Usage-Based Approach*. Cambridge: Cambridge University Press.
- Kaleta, Agnieszka 2014: *English Sentential Complementation: A Usage-Based Approach*. Piotrków Trybunalski: Naukowe Wydawnictwo Piotrkowskie.
- Langacker, Ronald W. 1987: *Foundations of Cognitive Grammar*. Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald W. 1991: *Foundations of Cognitive Grammar*. Vol. 2: *Descriptive Application*. Stanford: Stanford University Press.

- Langacker, Ronald W. 2000: *Grammar and Conceptualization*. Berlin, New York: Mouton de Gruyter.
- Langacker, Ronald W. 2002: *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Second edition. Berlin, New York: Mouton de Gruyter.
- Langacker, Ronald W. 2008: *Cognitive Grammar: A Basic Introduction*. Oxford, New York: Oxford University Press.
- Levshina, Natalia 2012: Comparing constructions: A usage-based analysis of the causative construction with *doen* in Netherlandic and Belgian Dutch. *Constructions and Frames* 4.1, 76–101.
- Noël, Dirk, Timothy Coleman 2010: Believe-type raising-to-object and raising-to-subject verbs in English and Dutch: A contrastive investigation in diachronic Construction Grammar. *International Journal of Corpus Linguistics* 15.2, 157–182.
- Rudzka-Ostyn, Brygida 1995: Metaphor, schema, invariance: The case of verbs of answering. In: Louis Goossens, Paul Pauwels, Brygida Rudzka-Ostyn, Anne-Marie Simon-Vandenberg, Johan Vanparys (eds.) 1995: *By Word of Mouth: Metaphor, Metonymy and Linguistic Action in a Cognitive Perspective*. Amsterdam, Philadelphia: John Benjamins, 205–244.
- Rudzka-Ostyn, Brygida 1996: The Polish dative. In: William van Belle, Willy van Langendonck (eds.) 1996: *The Dative*. Vol. 1: *Descriptive Studies*. Amsterdam, Philadelphia: John Benjamins, 341–394.
- Schmid, Hans-Jörg 2000: *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. Berlin, New York: Mouton de Gruyter.
- Stefanowitsch, Anatol, Stefan Thomas Gries 2003: Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8.2, 209–243.
- Stefanowitsch, Anatol, Stefan Thomas Gries 2005: Covarying collexemes. *Corpus Linguistics and Linguistic Theory* 1.1, 1–43.
- Szmrecsanyi, Benedikt 2010: The English genitive alternation in a cognitive sociolinguistics perspective. In: Dirk Geeraerts, Gitte Kristiansen, Yves Peirsman (eds.) 2010: *Advances in Cognitive Sociolinguistics*. Berlin, New York: Mouton de Gruyter, 141–165.
- Tomasello, Michael 2003: *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Verspoor, Marjolijn 2000: Iconicity in English complement constructions: Conceptual distance and cognitive processing levels. In: Kaoru Horie (ed) 2000: *Complementation: Interface of Cognitive, Functional and Typological Perspectives*. Amsterdam, Philadelphia: John Benjamins, 199–225.
- Wierzbicka, Anna 1988: *The Semantics of Grammar*. Amsterdam, Philadelphia: John Benjamins.
- Wulff, Stefanie 2006: Go–V vs. go–and–V in English: A case of constructional synonymy? In: Stefan Thomas Gries, Anatol Stefanowitsch (eds.) 2006, 101–126.

STRESZCZENIE

Metody korpusowe w semantyce kognitywnej na przykładzie angielskich konstrukcji dopełnieniowych

Niniejszy artykuł podejmuje zagadnienie przydatności metodologii korpusowej w badaniach nad semantyką struktur składniowych, prowadzonych w ramach językoznawstwa kognitywnego. Artykuł składa się z czterech części. Pierwsza z nich to *Wstęp*, który wprowadza podjętą tematykę, kreśli tło teoretyczne pracy, a w szczególności zwraca uwagę na aspekty kognitywnej koncepcji języka, które stanowią szczególne wyzwanie metodologiczne. Jednak zasadnicza część *Wstępu* dotyczy tych aspektów kognitywnej teorii języka, które wydają się kompatybilne z celami i metodami językoznawstwa korpusowego. Najważniejszą z nich jest teza o uzusie językowym, zgodnie z którą mentalne reprezentacje języka stanowią pochodną użycia języka, a ściślej mówiąc powstają na drodze uogólnień poczynionych na podstawie przypadków rzeczywistego użycia języka. Zatem korpusy językowe, które gromadzą znaczne zasoby rzeczywistych użyczeń języka, można uznać za naturalne źródło danych dla językoznawstwa kognitywnego.

Druga część artykułu zatytułowana *Językoznawstwo kognitywne spotyka językoznawstwo korpusowe: przegląd podejść korpusowych do zagadnień z semantyki kognitywnej* stanowi przegląd najważniejszych kierunków badawczych łączących metodologię korpusową z teoretyczną perspektywą językoznawstwa kognitywnego. Szczególną uwagę zwrócono na dwie wiodące metody badawcze: tzw. metodę profili behawioralnych (ang. behavioral profiles) oraz analizę kolostrukcyjną (ang. collocation analysis). Pierwsza z nich bazuje na szczegółowej anotacji pozyskanych z korpusu form językowych stanowiących przedmiot analizy. Anotacji podlegają zarówno aspekty formalno-składniowe, jak i semantyczne, a uzyskane na tej podstawie dane ilościowe poddawane są analizie statystycznej w celu uzyskania wglądu w typowe (statystycznie istotne) dla danej formy językowej tendencje dystrybucyjne. Metoda ta znalazła swe zastosowanie głównie w badaniach nad relacjami synonimii i polisemii leksykalnej. Analiza kolostrukcyjna z kolei została opracowana z myślą o badaniach nad semantycznymi aspektami konstrukcji składniowych. Opiera się ona na przekonaniu, że znaczenia form składniowych są zbieżne ze znaczeniami leksemów współwystępujących lub „wypełniających” dany schemat składniowy. Podejście to znalazło zastosowanie w badaniach nad semantyką schematycznych form składniowych (takich m.in. jak konstrukcja ditranzytywna), a w szczególności nad subtelnymi różnicami znaczeniowymi pomiędzy *quasi*-synonimicznymi konstrukcjami składniowymi.

Trzecia część niniejszego artykułu zatytułowana *Angielskie dopełnienie zdaniowe* ilustruje zakres i sposób użycia wymienionych wyżej metod na przykładzie badań nad anglojęzycznymi konstrukcjami dopełnieniowymi. Ta część artykułu rozpoczyna się od krótkiego przeglądu kognitywnej koncepcji konstrukcji dopełnieniowych (podpunkt 3.1).

W kolejnym podpunkcie (3.2) zaprezentowano przykładowe wyniki dystynktywnej analizy koleksemicznej (ang. distinctive collexeme analysis) dla *quasi*-synonimicznych konstrukcji bezokolicznikowych i gerundialnych występujących z czasownikami aspektualnymi i czasownikami emocji

(np. *start to Verb / start Verb ing; like to Verb / like Verb ing*). Analiza ta pozwoliła na określenie statystycznie istotnych różnic dystrybucyjnych pomiędzy analizowanymi konstrukcjami. Uzyskane wyniki, w połączeniu z dotychczasową wiedzą na temat procesów konceptualizacyjnych obecnych w języku, pozwoliły z kolei na wyciągnięcie wniosków dotyczących semantycznej motywacji w użyciu bezokolicznika i gerundium. Stwierdzono, że na najbardziej schematycznym (uogólnionym) poziomie reprezentacji semantycznej bezokolicznik z *to* jest motywowany schematem wyobraźniowym ŚCIEŻKI, gerundium natomiast pełni uogólnioną funkcję reifikacyjną (nominalizacyjną). Z kolei na poziomie bardziej szczegółowych reprezentacji, konstrukcje bezokolicznikowe z czasownikami aspektualnymi (np. *start to Verb*) przedstawiają opisywaną sytuację w kategoriach zmiany lub przejścia w nowy stan, natomiast odpowiadające im konstrukcje gerundialne (np. *start Verb ing*) desygnują początek czy też kontynuację ciągłego (dynamicznego) procesu.

W odniesieniu do czasowników emocji takich jak *like* czy *love* subtelna różnica znaczeniowa pomiędzy formami bezokolicznikowymi (np. *like to Verb*) a gerundialnymi (np. *like Verb ing*) sprowadza się do kontrastu pomiędzy pojęciem wolicjonalności (chęci zaistnienia danej sytuacji czy wydarzenia) a wyrażaniem pozytywnego (lub negatywnego) stosunku emocjonalnego do wydarzenia zdania dopełnieniowego. Zatem w tym drugim przypadku semantyka formy dopełnieniowej pozostaje zbieżna ze znaczeniem leksykalnym czasownika głównego; w przypadku bezokolicznika natomiast mamy do czynienia z dodatkowym elementem semantycznym, który stanowi pochodną użycia znacznika *to*. W artykule wykazano metaforyczne i metonimiczne związki łączące schematyczne reprezentacje semantyczne z ich bardziej uszczegółowionymi odpowiednikami.

W ostatnim podpunkcie tej części artykułu podano przykład zastosowania metody profili behawioralnych. Na przykładzie *promise* i *admit* wskazano na różnice znaczeniowe pomiędzy bezokolicznikowymi i gerundialnymi użyciami tych czasowników (*promise to Verb, admit Verb ing*) a ich odpowiednikami współwystępującymi ze zdaniami indykatywnymi z *that* (*promise that, admit that*). Dane dystrybucyjne pozyskane w toku przeprowadzonych badań pozwoliły na zidentyfikowanie czynnika emotywnego różnicującego użycie *promise to* i *promise that*. W przypadku konstrukcji z *admit* natomiast stwierdzono statystycznie istotny związek *admit Verb ing* z pojęciem przyznania się do winy (tj. do popełnienia przestępstwa lub wykroczenia), natomiast *admit that* z propozycjonalnym pojęciem akceptacji prawdziwości danego twierdzenia.

Przedstawione w punkcie trzecim przykłady zostały wykorzystane do sformułowania kilku ogólnych wniosków dotyczących przydatności, a także ograniczeń związanych z użyciem metod korpusowych w badaniach nad semantyką konstrukcji składniowych. Zostały one zaprezentowane w ostatnim punkcie artykułu zatytułowanym *Dyskusja i konkluzje*. Jak podkreślono, metody korpusowe umożliwiają wgląd w tendencje dystrybucyjne jednostek językowych. Ponadto pozwalają precyzyjnie określić częstotliwość występowania badanych konstrukcji, a tym samym umożliwiają odróżnienie prototypowych użyć od bardziej peryferyjnych konstrukcji. Szczególnie ważną rolę mają one do odegrania w badaniach nad *quasi*-synonimicznymi konstrukcjami, w przypadku których metody introspekcyjne często okazują się niewystarczające. Jednak pomimo swej przydatności, w zderzeniu z semantyką kognitywną metody korpusowe napotykają na istotne trudności. Jak zauważono, trudności te dotyczą głównie aspektu wyjaśniającego, a wynikają one głównie z wysokiego stopnia schematyczności (abstrakcyjności) znaczeń reprezentowanych przez kategorie składniowe oraz ich dynamicznego i elastycznego charakteru. Istotnym wyzwaniem dla metod ilościowych jest również często obserwowany brak bezpośrednich korelacji pomiędzy częstotliwościami korpusowymi a schematycznymi reprezentacjami językowymi. W końcowej konkluzji podkreślono, że pomimo swych ograniczeń metody korpusowe wnoszą istotny wkład w badania nad semantyką konstrukcji składniowych, choć badania te zazwyczaj są skazane na mniejszy stopień rygoru badawczego, niż ma to miejsce w przypadku empirycznych badań w tzw. „naukach twardych”.